

De invloed van evaluatie verklaard door de evaluatiesetting en de besluitvormingskenmerken van de organisatie

*Een casestudie naar de invloed van een benchmark in de
Nederlandse drinkwatersector (digitale versie)*

Open Universiteit Nederland
Faculteit Managementwetenschappen
Master of Science in Management

Naam: Jean-Pierre van Wissem
Studentnummer: 850954313
Datum: 7 juni 2012

Begeleider/examinator: Mw. M. Crijns
Medebeoordelaar: Dhr. G. Arendsen

SAMENVATTING

Centrale vraag van onderzoek

De laatste 40 jaar is veelvuldig onderzoek uitgevoerd naar de benutting van evaluaties. Hierdoor is het onderzoeksdomein 'evaluatie' verbreed en is een reeks factoren geïdentificeerd die invloed heeft op de benutting van evaluaties. Door een groeiend besef van het belang van de context waarin een evaluatie plaatsvindt zijn onderzoekers overgestapt van de 'benutting van evaluatie' naar 'evaluatie-invloed'.

Verschillende onderzoeken hebben aangetoond dat, wanneer deze evaluatie-invloed goed in kaart wordt gebracht, de invloed van evaluaties groter is dan menig manager voor mogelijk had gehouden.

Onduidelijk is echter welke factoren van belang zijn voor deze invloed. Veel organisaties hebben juist behoefte aan meer inzicht in de reden waarom de invloed van evaluaties per situatie verschillend kan zijn. Het identificeren van factoren die van belang zijn voor deze evaluatie-invloed is het vertrekpunt van deze scriptie.

Alle drinkwaterbedrijven in Nederland participeren jaarlijks in een gezamenlijke evaluatie (benchmark) om zo de doelmatigheid, kwaliteit en transparantie van de sector in beeld te brengen en te vergroten. Deze benchmark is in dit scriptieonderzoek gebruikt als casestudy met als centrale vraag: In welke mate verklaren de evaluatiesettingfactoren en de besluitvorming- en beleidsettingfactoren de aanwezigheid van evaluatie-invloed bij Waterleiding Maatschappij Limburg (WML) en Brabant Water?

In het onderzoek worden 2 subdoelstellingen onderscheiden, ten eerste het in kaart brengen van de aanwezige evaluatie-invloed als gevolg van de benchmark binnen de organisaties. Ten tweede het in kaart brengen van de factoren die een sleutelrol vervullen in de optimalisatie van de invloed van deze evaluatie. Een beter begrip van de factoren die van belang zijn voor evaluatie-invloed kan in de toekomst leiden tot betere evaluaties die gericht invloed kunnen uitoefenen.

Literatuurstudie en conceptueel model

Ten eerste is literatuurstudie verricht naar het begrip 'evaluatie-invloed', het ontstaan van evaluatie-invloed, de verschillende niveaus van evaluatie-invloed en welke verandermechanismen hieraan ten grondslag liggen. Op basis van de uitkomsten van de literatuurstudie is het theoretische kader geformuleerd. Dit omvat de wijze waarop verschillende factoren van invloed kunnen zijn op de benutting van evaluaties. Op basis van het theoretisch kader is vervolgens een conceptueel model ontwikkeld.

In het model worden verschillende factoren samengevat in 2 categorieën die van invloed zijn op de benutting van evaluaties: de evaluatiesetting en de besluitvorming- en beleidsetting. De categorie 'evaluatiesetting' kan verder onderverdeeld worden in een zestal factoren, namelijk: kwaliteit, geloofwaardigheid, relevantie, communicatie, resultaten en timing van de evaluatie. De categorie 'besluitvorming- en beleidsetting' kan eveneens verder onderverdeeld worden in een zestal factoren, namelijk: informatienoodzaak, besluitvormingskenmerken, politiek klimaat, concurrerende informatie, persoonlijke kenmerken en commitment naar de evaluatie toe. De twee hoofdcategorieën vormen het eerste deel van het conceptuele model dat in dit onderzoek gebruikt is. Vervolgens zijn hier de drie

verschillende niveaus van evaluatie-invloed aan toegevoegd, namelijk: individueel, interpersoonlijk en collectief.

Aan deze verschillende niveaus van evaluatie-invloed gaan verandermechanismen vooraf. Juist deze verandermechanismen stelt de evaluatie in staat om invloed uit te oefenen. De invloed beweegt zich tussen deze niveaus en dit kan dan ook gezien worden als een interactief proces. Elk van de niveaus correspondeert met een veranderproces dat is bestudeerd in de sociale- en gedragswetenschappen. De geïdentificeerde veranderprocessen zijn: generieke invloed, cognitief en affectief, motiverend en gedragsmatig. In dit onderzoek zijn alleen de veranderprocessen 'motiverend' en 'gedragsmatig' relevant.

Uitvoeren van onderzoek

Het is voor het onderzoek niet noodzakelijk om controle te hebben over het gedrag van participanten en de centrale onderzoeksvraag concentreert zich op de tegenwoordige tijd. De gevalsstudie is derhalve de meest voor de hand liggende onderzoeksmethode voor dit onderzoek. Vervolgens zijn twee van de vijf redenen om een gevalsstudie als enkelvoudig te benoemen van toepassing op deze casus. Dit onderzoek betreft dan ook een ingebedde enkelvoudige gevalsstudie. Dit onderzoek maakt gebruik van de dataverzamelmethode documentstudie en interviews. De interviews zijn semigestructureerd opgezet en uitgevoerd. De vragen zijn vooraf bepaald en vastgesteld, waarbij afhankelijk van de antwoorden ook verdiepende vragen in de routing van de vragenlijst zijn opgenomen. Aan deze interviews ligt een vragenlijst vanuit het theoretisch kader ten grondslag. Per deelvraag is beschreven welke methoden zijn gebruikt om data te verzamelen en hoe triangulatie heeft plaatsgevonden. Er zijn in totaal dertien belanghebbenden geïnterviewd; zeven respondenten bij Brabant Water en zes respondenten bij WML. In de digitale versie van dit onderzoek zijn de interviewverslagen wegens geheimhouding niet in de bijlagen opgenomen. De interviewverslagen zijn wel ter beschikking gesteld aan de faculteit managementwetenschappen van de OU.

Resultaten casestudy: Benchmark Nederlandse drinkwaterbedrijven

Evaluatiesettingfactoren en besluitvorming- en beleidsettingfactoren

In het geval van WML spelen met name de met elkaar samenhangende factoren 'kwaliteit', 'geloofwaardigheid' en 'resultaten' een rol van betekenis. Bij Brabant Water zijn dit eveneens de met elkaar samenhangende factoren 'kwaliteit', 'geloofwaardigheid' en 'resultaten' maar hieraan kunnen nog de factoren 'relevantie' en 'communicatie' worden toegevoegd. In het geval van WML zijn voornamelijk de factoren 'informatienoodzaak' en 'besluitvormingskenmerken' en in mindere mate de factoren 'politiek klimaat' en 'persoonlijke kenmerken' van belang voor de benchmark. Bij Brabant Water zijn dit dezelfde factoren, maar het belang van het politieke klimaat is groter.

Verschillende niveaus van evaluatie-invloed

In het geval van WML en Brabant Water worden alle drie de niveaus van evaluatie-invloed onderkend. Bij WML ligt de nadruk vooral op individuele en interpersoonlijke invloed, zoals het kritisch zijn op kosten en een meer externe focus. Belanghebbenden bij WML onderkennen ook collectieve invloed maar deze

is nauwelijks concreet en zeer beperkt. Bij Brabant Water ligt de nadruk op individuele maar vooral op collectieve invloed. De benchmark is een integraal onderdeel in de bedrijfsvoering van Brabant Water. Alle belangrijke beleidsmatige keuzes worden doorvertaald naar benchmarkresultaat en als zodanig ook meegenomen in de besluitvorming.

Invloed van de benchmark in relatie tot evaluatiesettingfactoren

De invloed van de benchmark bij WML wordt vooral verklaard door de factoren 'relevantie' en 'communicatie'. De benchmark wordt door verschillende belanghebbenden bij WML niet als relevant beschouwd. Dit wordt versterkt door het beeld dat belanghebbenden hebben dat de benchmark op hoofdlijnen wel betrouwbaar is maar op detailniveau niet. Belanghebbenden worden onvoldoende geïnformeerd betreffende de achtergrond en onderbouwing van de opgevoerde data in de benchmark. Dit kan direct herleid worden naar de factor 'communicatie'. De factoren 'geloofwaardigheid', 'kwaliteit' en 'resultaten' zijn sterk afhankelijk van de kwaliteit van invoer van gegevens. Bij een beperkte relevantie zal men sneller genoeg nemen met minder kwaliteit. Wanneer de benchmark een hogere relevantie heeft, zal de kwaliteit van de input en de hiermee samenhangende nauwkeurigheid en herkenbaarheid vanzelfsprekend groter zijn. De factor communicatie speelt hierin een rol. Binnen WML wordt maar beperkt aan een selecte groep medewerkers over de benchmark gecommuniceerd. Hierdoor zal de relevantie en invloed van de benchmark niet toenemen.

De invloed van de benchmark bij Brabant Water wordt vooral verklaard door de factoren 'relevantie', 'geloofwaardigheid' en 'communicatie'. Alle belanghebbenden bij Brabant Water betitelen de benchmark als zeer geloofwaardig. Ook hier zijn door belanghebbenden vraagtekens geplaatst op detailniveau en is de kwaliteit niet voldoende. Deze vraagtekens hebben echter niet betrekking op de eigen informatie maar meer op de opgevoerde datasets van andere bedrijven. In lijn met WML heeft ook dit gevolgen voor de resultaten en de geloofwaardigheid van de benchmark. De benchmark wordt daarnaast als zeer relevant gezien. De doelstelling om in de top 3 te staan voor wat betreft alle benchmarkprocessen is één op één gekoppeld aan de benchmark. Dit is een strategische doelstelling die door alle belanghebbenden gedragen en gecommuniceerd wordt. De benchmark wordt intern vooral ingezet als verbeterinstrument. Medewerkers worden uitgedaagd om beter en efficiënter te werken waarbij de benchmark als leidraad wordt gebruikt. In de communicatie wordt de benchmark vooral getypeerd als verbeterinstrument in plaats van vergelijking- of afrekeninstrument. Hierdoor verklaart ook de factor 'communicatie' een belangrijk aspect van de evaluatie-invloed van de benchmark bij Brabant Water.

Invloed van de benchmark in relatie tot de besluitvorming- en beleidsettingfactoren

De invloed van de benchmark bij WML wordt vooral verklaard door de factoren 'besluitvormingskenmerken', 'politiek klimaat', 'persoonlijke kenmerken' en 'commitment'. De besluitvorming binnen WML is zo divers beschreven dat niet gesproken kan worden van een eenduidig of gestructureerd proces. Door een lager draagvlak voor besluiten is de mogelijke invloed van de benchmark als gevolg van directiebesluiten laag. De benchmark sluit echter wel goed aan bij het interne politieke klimaat (kostenbeheersing). De invloed van de benchmark bij WML komt vooral tot stand door individuen die de benchmark begrijpen en waarderen. Ze zien het potentieel van de benchmark en gebruiken deze voor het verbeteren van hun

proces. De invloed vindt vooral plaats op afdelingsniveau en niet op bedrijfsniveau en is incidenteel van aard in plaats van structureel.

De invloed van de benchmark bij Brabant Water wordt vooral verklaard door de factoren 'besluitvormingskenmerken', 'politiek klimaat', 'persoonlijke kenmerken' en 'commitment'. De directie gelooft in de benchmark en heeft de benchmark tot een belangrijke doelstelling (top 3) gemaakt. Dit is in grote lijnen het gevolg van het commitment van enkele personen binnen Brabant Water naar de benchmark toe. Niet alleen committeert Brabant Water zich intern aan de benchmark, maar ook extern worden de doelstellingen gecommuniceerd en vastgelegd. Alhoewel het initiatief en de besluitvorming met name bij de directie ligt, wordt de benchmark toch enorm beleefd binnen de organisatie. De besluitvormingskenmerken hebben een belangrijke rol in dit proces. De collectieve invloed bij Brabant Water wordt met name veroorzaakt door de top 3 doelstelling. Veel besluiten die worden genomen zijn gebaseerd op deze doelstelling.

Conclusies en aanbevelingen

De resultaten van deze analyse laten een aantal grote overeenkomsten zien, maar ook een aantal duidelijke verschillen. Bij beide organisaties zijn alle drie de niveaus van invloed aanwezig maar is de collectieve invloed bij Brabant Water beduidend sterker aanwezig dan bij WML. Ook spelen zowel de factoren behorende tot de evaluatiesetting en de besluitvorming- en beleidsetting een rol van betekenis betreffende de aanwezigheid van evaluatie-invloed. Uit dit scriptieonderzoek blijkt dat de factoren behorende bij de besluitvorming- en beleidsetting van groter belang zijn.

In het theoretisch kader is ook het motiverende veranderproces opgenomen. Bij zowel Brabant Water als WML speelt dit veranderproces geen enkele rol van betekenis. Alle evaluatie-invloed wordt veroorzaakt door het gedragsmatige veranderproces. Verder blijkt ook uit de onderzoeksresultaten dat, conform het theoretisch kader, de evaluatie-invloed beweegt tussen de verschillende niveaus van invloed. De evaluatie-invloed bij Brabant Water is veel dynamischer, er vinden dus veel bewegingen plaats tussen de verschillende niveaus. Bij WML is dit duidelijk minder.

Een praktische implicatie die volgt uit dit onderzoek is dat de factoren 'commitment' en 'besluitvormingskenmerken' die behoren tot de categorie 'besluitvorming- en beleidsetting' en de factor 'relevantie' die behoort tot de categorie 'evaluatiesetting' van belang zijn voor evaluatie-invloed. Om een evaluatie effectief structurele invloed te kunnen laten uitoefenen binnen een organisatie zal men aandacht moeten besteden aan het commitment en de besluitvormingskenmerken van de organisatie. Al is de kwaliteit van de evaluatie nog zo goed, zonder voldoende draagvlak en structuur binnen de deelnemende organisaties zal dit niet leiden tot de gewenste structurele evaluatie-invloed.

Een andere implicatie die volgt uit dit onderzoek heeft betrekking op de onderzochte verandermechanismen. Bij beide organisaties is de evaluatie-invloed vooral tot stand gekomen als gevolg van de gedragsmatige verandermechanismen. De motivationele verandermechanismen spelen nauwelijks een rol van betekenis. Indien de gedragsverandering laag is, of lager dan gewenst kan men, om de evaluatie-

invloed te vergroten – de aandacht wellicht focussen op datgene wat belanghebbenden triggert om hun gedrag te veranderen.

Er is een duidelijke verband tussen de factoren 'communicatie' en 'relevantie'. Uit de onderzoeksresultaten blijkt dat geregeld communiceren betreffende de opzet, inhoud en resultaten van de evaluatie, positief werkt op de relevantie van de evaluatie. Een evaluatie kan dus, door regelmatig en doelgericht te communiceren, eenvoudig relevanter gemaakt worden binnen een organisatie.

INHOUDSOPGAVE

1	ONDERWERP, RELEVANTIE EN DOELSTELLING	10
1.1	INLEIDING	10
1.2	MAATSCHAPPELIJKE RELEVANTIE.....	10
1.3	PERSOONLIJKE RELEVANTIE MET BETREKKING TOT HET ONDERWERP	11
1.4	WETENSCHAPPELIJKE RELEVANTIE VAN HET ONDERWERP	11
1.5	PROBLEEMSTELLING	12
1.5.1	<i>Doelstelling onderzoek</i>	<i>12</i>
1.5.2	<i>Centrale vraag en deelvragen</i>	<i>13</i>
2	THEORETISCH KADER	14
2.1	EVALUATIE-INVLOED	14
2.1.1	<i>Van benutting van evaluaties naar evaluatie-invloed.....</i>	<i>14</i>
2.1.2	<i>Kirkhart's Integrated Theory of Influence.....</i>	<i>15</i>
2.1.3	<i>Mark en Henry Three Level Model.....</i>	<i>17</i>
2.1.4	<i>Mark en Henry Comprehensive Theory of Evaluation Influence.....</i>	<i>19</i>
2.2	FACTOREN DIE DE BENUTTING VAN EVALUATIES BEÏNVLOEDEN	22
2.2.1	<i>Overzicht van relevante onderzoeken en literatuur</i>	<i>22</i>
2.2.2	<i>Onderzoek Bober & Bartlett (2004).....</i>	<i>24</i>
2.3	CONCEPTUEEL MODEL	25
3	METHODOLOGIE.....	27
3.1	KEUZE VAN DE ONDERZOEKSSTRATEGIE	27
3.1.1	<i>Holistisch of ingebedde enkelvoudige gevalsstudie?.....</i>	<i>28</i>
3.1.2	<i>Keuze van de gevalsstudie.....</i>	<i>28</i>
3.1.3	<i>Keuze van de analyse-eenheden.....</i>	<i>29</i>
3.1.4	<i>Keuze van de waarnemingseenheden</i>	<i>29</i>
3.2	DATAVERZAMELING	29
3.2.1	<i>De methoden bij deelvraag 1, 2, 3 en 4.....</i>	<i>31</i>
3.2.2	<i>De methoden bij deelvraag 5 en 6.....</i>	<i>31</i>
3.2.3	<i>De methoden bij deelvraag 7 en 8.....</i>	<i>31</i>
3.2.4	<i>De methoden bij deelvraag 9, 10 en 11.....</i>	<i>31</i>
3.3	BEPERKINGEN VAN DE ONDERZOEKSRISULTATEN.....	32
4	RESULTATEN VAN HET ONDERZOEK.....	34
4.1	EVALUATIESETTINGFACTOREN	34
4.2	BESLUITVORMING EN BELEIDSETTINGFACTOREN	36
4.3	NIVEAUS VAN INVLOED	37
4.3.1	<i>Niveaus van invloed in het geval van WML.....</i>	<i>38</i>

4.3.2	<i>Niveaus van invloed in het geval van Brabant Water</i>	39
4.3.3	<i>Verschillen en overeenkomsten WML en Brabant Water</i>	40
4.4	TYPEN VERANDERMECHANISMEN	40
4.5	INVLOED FACTOREN EVALUATIESETTING	41
4.5.1	<i>Invloed factoren evaluatiesetting in het geval van WML</i>	41
4.5.2	<i>Invloed factoren evaluatiesetting in het geval van Brabant Water</i>	42
4.6	INVLOED FACTOREN BESLUITVORMING- EN BELEIDSETTING	43
4.6.1	<i>Invloed factoren besluitvorming- en beleidsetting in het geval van WML</i>	43
4.6.2	<i>Invloed factoren besluitvorming- en beleidsetting in het geval van Brabant Water</i>	44
4.7	VERKLARING VERSCHIL VAN INVLOED	46
5	CONCLUSIES, DISCUSSIE EN AANBEVELINGEN	47
5.1	CONCLUSIES VAN HET ONDERZOEK	47
5.2	DISCUSSIE NAAR AANLEIDING VAN HET ONDERZOEK	51
5.2.1	<i>Analyseren onderzoeksresultaten ten opzichte van theoretisch kader</i>	51
5.2.2	<i>Bijdragen onderzoeksresultaten aan wetenschappelijke kennis en literatuur</i>	51
5.2.3	<i>Betrouwbaarheid en validiteit van het onderzoek</i>	52
5.3	AANBEVELINGEN NAAR AANLEIDING VAN HET ONDERZOEK	53
5.3.1	<i>Praktische implicaties volgend uit het onderzoek</i>	53
5.3.2	<i>Aanleiding voor vervolgonderzoek</i>	54
6.	LITERATUUR	56
6.1	WETENSCHAPPELIJKE LITERATUUR	56
6.2	DOCUMENTEN	59
BIJLAGE 1	VRAGENLIJST FACTOREN INTERVIEW BELANGHEBBENDEN	60
BIJLAGE 2	VRAGENLIJST INVLOED INTERVIEW BELANGHEBBENDEN	62
BIJLAGE 3	INTERVIEWVERSLAGEN BELANGHEBBENDEN	63

1 ONDERWERP, RELEVANTIE EN DOELSTELLING

In dit hoofdstuk wordt ten eerste het onderzoeksonderwerp toegelicht. Vervolgens wordt ingegaan op de maatschappelijke relevantie van - en mijn persoonlijke betrokkenheid bij - het onderwerp. Tevens worden de thema's 'evaluatie-invloed' en 'factoren van invloed' in een groter geheel geplaatst. Dit gebeurt door een (korte) plaatsbepaling van het onderwerp binnen de wetenschappelijke literatuur. Tot slot worden de doelstelling en de onderzoeksvragen toegelicht.

1.1 Inleiding

Alle drinkwaterbedrijven in Nederland participeren jaarlijks in een gezamenlijke evaluatie om zo de doelmatigheid, kwaliteit en transparantie van de sector in beeld te brengen en te vergroten. In deze evaluatie, genaamd benchmark 'Water in Zicht', komen vier thema's aan de orde, namelijk: waterkwaliteit, dienstverlening, milieu en financiën & efficiency. Sinds 2011 is deelname aan deze benchmark verplicht als onderdeel van de nieuwe drinkwaterwet.

Deze benchmark is in dit scriptieonderzoek gebruikt als casestudy om een meer fundamenteel onderwerp aan de orde te stellen, namelijk: de invloed van evaluatie en de mate waarin de evaluatie-invloed verklaard kan worden door factoren die toebedeeld worden aan de evaluatiesetting en factoren die behoren tot de besluitvorming- en beleidsetting. Het object van dit onderzoek zijn de belanghebbenden van de benchmark binnen Waterleiding Maatschappij Limburg (WML) en Brabant Water (BW).

Het onderwerp is uitgewerkt aan de hand van de verschillende niveaus van evaluatie-invloed die in de wetenschappelijke literatuur onderscheiden worden. Hierbij is ook gekeken naar verandermechanismen, die worden toegepast door belanghebbenden, om te komen tot deze invloed.

Scriptieonderwerp:

'In welke mate verklaren de evaluatiesettingfactoren en de besluitvorming- en beleidsettingfactoren de aanwezigheid van evaluatie-invloed bij WML en Brabant Water?'

1.2 Maatschappelijke relevantie

In 1975 hebben aanpassingen in de toenmalige drinkwaterwet geleid tot een fusiegolf die de drinkwatersector structureel heeft veranderd. In die tijd kende Nederland meer dan 100 drinkwaterbedrijven verdeeld over verschillende provincies en gemeenten. Door de wetsbepaling dat een drinkwaterbedrijf over minimaal 100.000 aansluitingen moest beschikken, heeft er een verandering plaatsgevonden van kleine lokale bedrijven naar grotere bedrijven in de vorm van naamloze vennootschappen. De aandelen kwamen in handen van verschillende overheidsinstanties zoals provincies en gemeenten (Blokland, Braadbaart, & Schwartz, 1999). Deze wetsbepaling en de verwachte liberalisering van de Europese markten zorgden ervoor dat er in 1985 nog 91 drinkwaterbedrijven binnen Nederland opereerden, in het jaar 2000 waren dat er nog 23 en in 2011 nog maar 10.

In 1996 heeft in opdracht van het Ministerie van Economische Zaken een studie plaatsgevonden waarin is onderzocht of drinkwater geleverd kon worden tegen een aanzienlijk lagere prijs. Een kostenreductie zou dan gerealiseerd worden door het privatiseren van de drinkwaterbedrijven. Hierop heeft de Vereniging van Waterbedrijven in Nederland (Vewin) actie ondernomen en is destijds de eerste nationale benchmark 'Water in Zicht' uitgevoerd. Met deze benchmark werd aangetoond dat drinkwaterbedrijven transparant opereerden en de geleverde prestaties - en de hieruit voortvloeiende drinkwatertarieven - billijk waren.

De laatste twee decennia heeft er tevens een cultuuromslag plaatsgevonden die zich kenmerkt door een strategische marktbenadering die meer verwantschap heeft met de profit- dan de non-profit sector. Benchmarking is hiervan een onderdeel (Dattakumar & Jagadeesh, 2003, pp. 176-209); (Dorsch & Yasin, 1998, pp. 91-115).

Een acceptabel drinkwatertarief is van groot belang voor de Nederlandse samenleving. Water is één van de belangrijkste grondstoffen op onze planeet en dient voor een ieder betaalbaar te zijn. Dit kan beschouwd worden als een belangrijke stap richting 'social betterment'. Social betterment wordt getypeerd als de belangrijkste doelstelling van evaluaties (Mark & Henry, 2003, pp. 295-296). Het optimaliseren van de invloed van de benchmark is derhalve een volgende stap in het uitvoeringsproces van social betterment.

1.3 Persoonlijke relevantie met betrekking tot het onderwerp

Momenteel ben ik ruim 3 jaar als bedrijfscontroller werkzaam bij Waterleiding Maatschappij Limburg (WML). De organisatie WML is onderverdeeld in 4 sectoren: Markt & Infrastructuur, Klant & Operatie, Bedrijfsondersteuning en Directie. Als bedrijfscontroller ben ik verantwoordelijk voor de Sector Markt & Infrastructuur en daarnaast ben ik coördinator van de benchmark. De resultaten van de benchmark vormen de basis voor verdere analyse en de vertaalslag naar actiepunten en beleid. Door inzicht in aspecten waarop collega-bedrijven beter scoren dan WML kan er intern gewerkt worden aan verbeterpunten. Via dit onderzoek kan ik bijdragen aan de optimalisatie van het gebruik van de benchmark en kan deze kennis gedeeld worden met interne en externe partners.

In mijn scriptie wil ik de evaluatie-invloed bij WML en Brabant Water in kaart brengen en onderzoeken welke invloed volgens belanghebbenden kan worden toegewezen aan de verschillende factoren. Met andere woorden, welke factoren zorgen ervoor dat de evaluatie invloed heeft? Op basis van de onderzoeksresultaten kunnen we de effectiviteit en invloed van de benchmark in de nabije toekomst vergroten. Dat is mijn persoonlijke drijfveer voor de keuze van dit onderzoeksonderwerp.

1.4 Wetenschappelijke relevantie van het onderwerp

De laatste 40 jaar is veelvuldig onderzoek uitgevoerd naar de 'de benutting van evaluaties'. Hierdoor is het onderzoeksdomein 'evaluatie' verbreed en is een reeks factoren geïdentificeerd die invloed heeft op de benutting van evaluaties (Alkin, 1985, pp. 14-15). Hoewel een aantal onderzoekers zich aanvankelijk heeft gefocust op het concept 'potentie van de benutting' (Cousins & Leithwood, 1986, pp. 331-364) concentreerden de meeste onderzoekers zich op de verschillende aspecten van evaluatietoepassingen.

Door een groeiend besef van het belang van de context waarin een evaluatie plaatsvindt zijn onderzoekers overgestapt van de 'benutting van evaluatie' naar 'evaluatie-invloed' (Kirkhart, 2000, pp. 6-7). Volgens Kirkhart kan evaluatie-invloed verder onderverdeeld worden in drie dimensies, te weten: bron, intentie en tijd. Mark & Henry (2003, pp. 297-298) hebben op basis van Kirkhart's theorie een model ontwikkeld waarbij invloed uitgeoefend kan worden op drie niveaus, namelijk: individueel, interpersoonlijk en collectief. Beide modellen kunnen gebruikt worden om toekomstige evaluaties verder te verfijnen door te focussen op die dimensie of dat niveau waarmee de meeste invloed uitgeoefend kan worden.

Een andere richting van aanvullend onderzoek concentreert zich op het ontwikkelen van evaluatie-invloedtheorieën, onderzoek naar veranderprocessen en hoe de invloed van evaluaties verbeterd kan worden (Cummings, 2002, pp. 1-7), (Mark & Henry, 2003, pp. 293-314), (Mark & Henry, 2004, pp. 35-57), (Weiss, 1998, pp. 21-33). Op basis van dit type onderzoek is het mogelijk om een extra stap te zetten richting 'social betterment'. Mark & Henry (2004, pp. 45-53) hebben hieraan een belangrijke bijdrage geleverd met hun 'Comprehensive Theory of Evaluation Influence'. Op basis van deze theorie wordt duidelijk gemaakt hoe groot het belang is van de onderliggende verandermechanismen binnen een evaluatie. Verschillende onderzoeken hebben aangetoond dat, wanneer deze evaluatie-invloed goed in kaart wordt gebracht, de invloed van evaluaties groter is dan menig manager voor mogelijk had gehouden. Onduidelijk is echter welke factoren van belang zijn voor deze veranderprocessen.

De variëteit in achtergronden geeft het belang aan dat onderzoekers toekennen aan contextuele achtergronden en hoe deze de invloed van evaluaties kan bepalen (Shulha & Cousins, 1997, p. 204). Hoewel er veel is onderzocht, bestaan er ook nog veel vragen omtrent de benutting van de resultaten van evaluaties en de factoren die hierbij een rol spelen. Veel organisaties hebben juist behoefte aan meer inzicht in de redenen waarom de invloed van evaluaties per situatie verschillend kan zijn.

Het identificeren van de factoren die van belang zijn voor de hiervoor beschreven veranderprocessen vormt het vertrekpunt van deze scriptie. De wetenschappelijke relevantie van dit onderzoek is tweeledig. Ten eerste kan dit onderzoek de betrouwbaarheid van de door Mark en Henry (2004, p. 47) geïdentificeerde niveaus van evaluatie-invloed en bijbehorende verandermechanismen vergroten door empirisch te onderzoeken of deze daadwerkelijk in de praktijk worden aangetroffen. Ten tweede maakt dit onderzoek inzichtelijk welke factoren volgens de belanghebbenden bepalend zijn voor de aanwezigheid van evaluatie-invloed. Bober en Bartlett (2004, pp. 363-383) hebben dit onderzocht voor een viertal Universiteiten. Als basis gebruikten zij het model van Cousins en Leithwood (1986, pp. 331-364) en kwamen tot de conclusie dat binnen de onderzochte Universiteiten de evaluatiesettingfactoren van veel groter belang waren voor de benutting van evaluaties dan de besluitvorming- en beleidsettingfactoren. Dit onderzoek is een uitbreiding op de bestaande theorie en tracht in kaart te brengen of de conclusies van Bober en Bartlett ook geldend zijn binnen de Nederlandse drinkwatersector.

1.5 Probleemstelling

1.5.1 Doelstelling onderzoek

Het doel van dit onderzoek is het in kaart brengen van de evaluatie-invloed bij WML en Brabant Water

en welke invloed volgens belanghebbenden kan worden toegewezen aan de verschillende factoren. En meer in het bijzonder: welke verandermechanismen hiervoor verantwoordelijk zijn geweest. Op basis hiervan wordt in kaart gebracht welke factoren van belang zijn voor het hebben van evaluatie-invloed en of deze factoren behoren tot de evaluatiesetting of de besluitvorming- en beleidsetting.

Bij de uitvoering van het onderzoek is de benchmark 'Water in Zicht' binnen de Nederlandse drinkwatersector als casestudy opgenomen. In het onderzoek worden 2 subdoelstellingen onderscheiden, namelijk:

- 1 Het in kaart brengen van de belanghebbenden binnen de organisatie en hun invloed op het evaluatieproces ten aanzien van het gebruik van de benchmark;
- 2 Het in kaart brengen van de factoren die een sleutelrol vervullen in de optimalisatie van de invloed van deze evaluatie.

1.5.2 Centrale vraag en deelvragen

De navolgende centrale onderzoeksvraag ligt ten grondslag aan het onderzoek.

In welke mate verklaren de evaluatiesettingfactoren en de besluitvorming- en beleidsettingfactoren de aanwezigheid van evaluatie-invloed bij WML en Brabant Water?

Om de centrale onderzoeksvraag te beantwoorden, is deze vraag opgedeeld in de volgende onderliggende elf deelvragen.

1. Wat wordt in wetenschappelijk onderzoek verstaan onder evaluatie-invloed?
2. Welke verschillende niveaus van evaluatie-invloed kunnen worden onderscheiden?
3. Welke verandermechanismen kunnen worden onderscheiden die de totstandkoming van evaluatie-invloed beïnvloeden?
4. Welke factoren kunnen de benutting van evaluaties beïnvloeden?
5. Welke factoren in het kader van de evaluatiesetting van de benchmark spelen een rol in het geval van WML en Brabant Water?
6. Welke factoren in het kader van de besluitvorming- en beleidsetting van de benchmark spelen een rol in het geval van WML en Brabant Water?
7. Welke verschillende niveaus van evaluatie-invloed worden, als gevolg van de benchmark, onderkend door de belanghebbenden in het geval van WML en Brabant Water? Op basis van welke praktijkvoorbeelden kan verklaard worden waaruit deze invloed bestaat?
8. Welke typen verandermechanismen liggen aan de basis van de totstandkoming van deze evaluatie-invloed in het geval van WML en Brabant Water?
9. Op welke manier verklaren de factoren in het kader van de evaluatiesetting van de benchmark de invloed van de benchmark in het geval van WML en Brabant Water?
10. Op welke manier verklaren de factoren in het kader van de besluitvorming- en beleidsetting van de benchmark de invloed van de benchmark in het geval van WML en Brabant Water?
11. Op welke manier kan het verschil in invloed van de evaluatie verklaard worden door de evaluatiesetting- of besluitvorming- en beleidsettingfactoren?

2 THEORETISCH KADER

In dit hoofdstuk wordt het theoretisch kader toegelicht. Op basis van literatuuronderzoek zijn de belangrijkste theoretische modellen in kaart gebracht en wordt het conceptuele model toegelicht.

2.1 Evaluatie-invloed

In deze paragraaf wordt het theoretische kader beschreven van het begrip 'evaluatie-invloed'. Hiermee wordt antwoord gegeven op de eerste en tweede deelvraag van het onderzoek. Ten eerste wordt de betekenis en het ontstaan van evaluatie-invloed beschreven. Vervolgens worden de verschillende niveaus van evaluatie-invloed toegelicht. Ten slotte worden de verandermechanismen behorend bij de verschillende niveaus benoemd.

2.1.1 Van benutting van evaluaties naar evaluatie-invloed

Binnen het onderzoeksgebied 'evaluatie' wordt al bijna 40 jaar gesproken en gediscussieerd over het onderwerp 'toepassen of benutting van evaluatie' ('use or utilization'). Mark & Henry (2003, p. 294) leggen uit dat de benutting van evaluatie de meeste aandacht heeft gekregen midden jaren '70 tot begin jaren '80 van het vorige decennium. Hernieuwde belangstelling is weer ontstaan eind jaren '90 tot de beginjaren van het huidige decennium. Het gezamenlijke comité voor standaarden voor educatieve evaluaties (1994) heeft het toepassen van evaluaties opgenomen in de lijst van de belangrijkste standaardisatiecategorieën. Onder deze standaarden hebben Lawrence, Gullickson en Toal (2007, pp. 275-276) er 4 geïdentificeerd die van belang zijn voor de benutting van evaluaties, te weten:

1. Stakeholder identificatie;
2. Informatiescope en selectie;
3. Duidelijkheid van rapportage;
4. Evaluatie-impact.

De afgelopen decennia zijn er verschillende definities van de benutting van evaluaties en/of type benutting van evaluaties tot stand gekomen. Het woord 'benutting' is vaak onderwerp van discussie tussen de beoordelaar en cliënt bij de bespreking van de voordelen, het belang en de impact van de evaluatie. Veel van de beschikbare literatuur over evaluaties concentreert zich dan ook op het identificeren van de verschillende factoren die de benutting van evaluaties beïnvloeden. Alkin et al (1979) identificeren 50 factoren, Huberman en Gather Thurler (1991) identificeren 47 factoren die de benutting van evaluaties beïnvloeden. Uiteindelijk stelt Balthasar (2006, p. 357) op basis van de 'shoppinglist' van Shulha en Cousins (1997, pp. 195-208) en andere literatuur een short list samen van 3 categorieën die de benutting van evaluaties beïnvloeden, te weten: de institutionele context, de omgeving en het proces van evalueren. Shadish, Cook en Leviton (1991, pp. 36-67) benadrukken het belang van benutting aangezien dit een belangrijk thema is binnen de evaluatietheorie. Zij stellen verder dat benutting essentieel is voor de verdere legitimatie van het onderzoeksthema 'evaluatie'. Alhoewel reeds veel van deze factoren zijn onderzocht en geïdentificeerd, zijn onderzoekers het erover eens dat de benutting van evaluaties verder aandacht en onderzoek nodig heeft (Lawrenz, Huffman, & McGinnis, 2007); (Preskill & Caracelli, 1997, pp. 209-225).

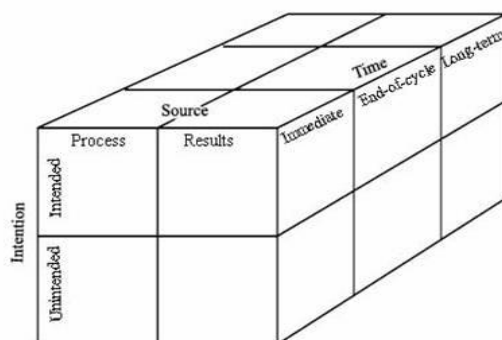
Het ontwikkelen van een definitie van de benutting van evaluaties was van groot belang voor de beoordelaars. De belangrijkste doelstelling van een evaluatie is het benutten van de resultaten. Zonder een eenduidige definitie wordt het voor de beoordelaars erg moeilijk om in te schatten hoe de resultaten van de evaluatie worden benut en wat de reikwijdte is geweest van de evaluatie. In hun review van de literatuur vermelden Cousins en Leithwood (1986) enkele onderzoeken die focussen op het concept 'potentie van het gebruik'. De meerderheid van de onderzoeken concentreerde zich echter meer op de concrete aspecten van de benutting van evaluatie. Hierbij kan men denken aan benutting ten behoeve van besluitvorming of educatieve doeleinden.

In de eerste onderzoeken naar de benutting van evaluaties werd benutting puur gezien vanuit het perspectief van het besluitvormingsproces. Dit perspectief staat nu bekend als instrumentele benutting (Preskill & Torres, 1996). Vervolgens is de definitie van benutting steeds verder uitgebreid en zijn hier onderdelen aan toegevoegd, zoals de conceptuele benutting, de politieke/overredende/symbolische benutting en de procesbenutting (Shulha & Cousins, 1997, p. 199). In de meer recente literatuur over de benutting van evaluaties is nu ook het misbruik van evaluaties opgenomen.

De *benutting* van evaluaties wordt tegenwoordig steeds vaker vervangen door de *invloed* van evaluaties aangezien deze term beter de veranderingen verwoord die optreden als gevolg van een evaluatie (Christie, 2007, pp. 8-25); (Mark & Henry, 2003); (Kirkhart, 2000). Het onderzoeksgebied 'evaluatie' heeft zich dus verbreed door niet meer alleen te focussen op de benutting maar het accent te verschuiven naar de invloed van evaluaties. Het woord 'use' wordt vervangen door 'Influence'. Kirkhart (2000, p. 7) prefereert de term invloed omdat deze term meer omvat dan enkel het gebruik. Het gaat dan om: "the capacity or power of persons or things to produce effects on others by intangible or indirect means".

2.1.2 Kirkhart's Integrated Theory of Influence

Volgens Kirkhart focust het concept 'evaluatiebenutting' te veel op resultaten. Het ontbreekt hierbij aan een goed theoretisch kader waardoor het begrip 'benutting' erg eendimensionaal en episodisch is. Om de invloed van evaluaties te kunnen bepalen dient niet alleen gekeken te worden naar de individuele componenten maar naar het totaal. Kirkhart heeft hiervoor reeds in 1995 een model ontwikkeld (zie figuur 1).



Figuur 1. Kirkhart's Integrated Theory of Influence

Dit model is in 2000 aangepast en heet 'de integrale theorie van evaluatie-invloed'. In deze theorie worden drie dimensies van invloed onderkend: de bron, tijd en intentie. Deze drie dimensies kunnen nog verder onderverdeeld worden in vier categorieën, namelijk: ongewilde procesgebaseerde invloed, gewilde procesgebaseerde invloed, ongewilde resultaatgebaseerde invloed en gewilde resultaatgebaseerde invloed.

De bron

De bron is voor het eerst beschreven door Henry en Rog (1998) en verwijst naar de referentie (evaluatierapport) waar de invloed van afstamt. De evaluatie-invloed wordt gerepresenteerd door de impact die een evaluatie heeft op personen/individuen door zowel immateriële als indirecte middelen. Sommige immateriële middelen zoals instrumentele, conceptuele en persuasieve evaluaties kunnen gezien worden als resultaatgebaseerde invloed. Aan de andere kant is er procesinvloed die ontstaat door deelname aan een evaluatieproces ongeacht de uitkomst. Invloeden kunnen proces- of resultaat gebaseerd zijn. Historisch gezien is de invloed van evaluaties voornamelijk gezien vanuit het kader van de benutting van de resultaten (Johnson, 1998, pp. 93-110); (Shulha & Cousins, 1997, pp. 195-208). Gelijktijdig bestond er wetenschappelijke literatuur met betrekking tot de interpersoonlijke invloeden van het evaluatieproces. Pas na de ontdekking van procesbenutting door Patton (1998, pp. 225-233) zijn deze twee stromingen samengevoegd.

Resultaatgebaseerde benutting wordt altijd bekeken vanuit instrumentele, conceptuele of symbolische benutting. Symbolische benutting wordt door anderen ook gezien als politieke en/of legislatieve benutting. Procesgerelateerde invloed wordt veroorzaakt door het proces van evalueren en niet door de resultaten van de evaluatie. Procesbenutting komt voor het eerst voor in de literatuur als ondersteuning van een onderzoek naar resultaatbenutting. Pas na verschillende onderzoeken die herhaaldelijk procesbenutting onderzochten, onafhankelijk van resultaatbenutting (Whitmore, 1991, pp. 1-7), is gebleken dat procesbenutting een belangrijke component is. Greene (1988b, pp. 91-116) onderkent hierbij drie verschillende dimensies binnen de procesinvloed, namelijk: cognitief, affectief en politiek.

De intentie

De intentie verwijst naar de betrokkenheid van de beoordelaar, cliënt en stakeholder in het plannen en sturen van de invloed die gegenereerd wordt door de evaluatie. De intentie leidt tot bedoelde invloed en verwijst naar een directe link tussen de intentie van de gebruikers en de geëffectueerde invloed van de evaluatie. Onbedoelde invloed refereert naar niet geplande invloed en consequenties als gevolg van de evaluatie. Zoals eerder beschreven kunnen zowel resultaat- als procesgebaseerde invloed intentioneel of niet intentioneel plaatsvinden. De intentie om te beïnvloeden markeert bij andere gebruikerstheorieën de grens tussen gebruik en misbruik van evaluaties (Alkin, 1990); (Alkin, Daillak, & White, 1979).

De tijd

De tijd als dimensie refereert naar het belang om te kunnen kijken naar het moment waarop de invloed van een evaluatie is geëffectueerd. Dit effect kan direct optreden, aan het einde van de evaluatiecyclus of op lange termijn. Elk van deze typen kan weer gecategoriseerd worden in een van de eerder gemelde dimensies (bron en intentie). Deze dimensie benadrukt ook de dynamische natuur van invloed. Invloeden die direct meetbaar zijn vinden vaak plaats gedurende het evaluatieproces en zijn dus procesgerelateerd,

een voorbeeld hiervan is het verder ontwikkelen van bestaande en nieuwe competenties door medewerkers als gevolg van deelneming aan de evaluatie.

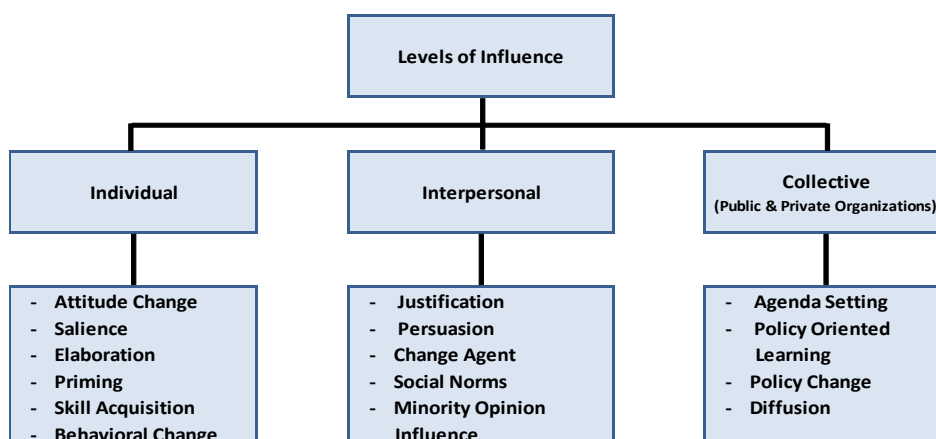
Er bestaan ook directe resultaatgebaseerde invloeden. Tijdens een evaluatie wordt na de eerste data-verzamelingsronde geconstateerd dat de opgeleverde data veel meer informatie bevat dan waarop men had geanticipeerd. Op basis hiervan wordt besloten om additioneel budget beschikbaar te stellen voor verder onderzoek. Deze korte beschrijving is een voorbeeld van directe invloed als gevolg van resultaten. Invloed aan het einde van de evaluatiecyclus kan zowel op basis van proces als resultaten en zowel intentioneel als niet intentioneel zijn. Lange termijn invloeden worden vaak pas lang nadat de evaluatie heeft plaatsgevonden herkend. De beoordelaar zal dus regelmatig, zeker nadat het eindrapport van de evaluatie is gepresenteerd, moeten controleren of er nog gewenste/ongewenste invloeden hebben plaats gevonden als gevolg van de evaluatie. Indien dit niet wordt gedaan zullen veel invloeden ook niet geïdentificeerd kunnen worden.

Kirkhart (2000, pp. 18-19) benoemt 9 mogelijke toepassingen:

1. Ter verduidelijking van de discussies over gebruik;
2. Het in kaart kunnen brengen van de invloed van een specifieke evaluatie;
3. Opvolgen van de evoluerende patronen van invloed over langere periodes;
4. Onderzoeken van misbruik en gebruik;
5. Verbeteren van de validiteit van onderzoeken naar invloed van evaluaties;
6. Faciliteren van metaonderzoeken naar invloed;
7. Opvolgen evolutie van de evaluatietheorie;
8. Vergelijken van evaluatietheorieën;
9. Ondersteunen bij het construeren van nieuwe theoretische raamwerken betreffende evaluatie.

2.1.3 Mark en Henry Three Level Model

Er is in beperkte mate wetenschappelijke literatuur beschikbaar over de invloed van evaluaties. Een uitzondering zijn Mark en Henry (2003), zij hebben een model ontwikkeld ter verbetering van het begrip van de invloed van evaluaties, zie figuur 2. Zij classificeren veranderprocessen en uitkomsten die beïnvloed kunnen worden door evaluaties op drie verschillende niveaus, te weten: individueel, interpersoonlijk en collectief. Deze drie niveaus corresponderen met veranderprocessen die zijn bestudeerd in de sociale- en gedragswetenschappen. Ter illustratie is het model van Mark en Henry toegevoegd (figuur 2), gevolgd door een beschrijving van elk niveau van invloed en de respectievelijke verandermechanismen.



Figuur 2. Mark en Henry's Three Level Model

Individueel

Op het individuele niveau benoemen Mark en Henry zes mogelijke uitkomsten, ook wel verandermechanismen genoemd, te weten:

- 1 Verandering van houding (herzien en veranderen van mening);
- 2 Saillantie (inschatten van het belang van het probleem);
- 3 Uitwerking (toegenomen diepgang bij het verwerken van specifieke problemen);
- 4 Priming (onder de aandacht brengen van de juiste personen);
- 5 Verwerven van vaardigheden (evaluatie competentie);
- 6 Verandering van gedrag (nieuwe concepten kunnen er toe leiden dat personen hun gedrag gaan veranderen).

Al deze verandermechanismen kunnen gezien worden als een aparte/individuele kubus in Kirkhart-model (2000, p. 8).

Interpersoonlijk

Op het interpersoonlijke niveau van evaluatie-invloed wordt gerefereerd naar activiteiten/stappen die personen kunnen ondernemen om de houding en gedrag van andere personen te kunnen wijzigen en/of beïnvloeden. Mark en Henry identificeren vijf mechanismen die deze veranderingen in houding en gedrag bij anderen te weeg kunnen brengen, te weten:

- 1 Rechtvaardiging (resultaten gebruiken om bestaande positie/houding te rechtvaardigen);
- 2 Overreding (de resultaten gebruiken om andere personen te overtuigen);
- 3 Veranderinstrument (resultaten gebruiken om anderen aan te zetten tot verandering);
- 4 Sociale normen (resultaten leiden tot het bijstellen van een normenkader)
- 5 Invloed van de opinie van een minderheid (de opinie van een kleine groep overheerst).

Collectief

Het collectieve niveau refereert naar de invloed van de evaluatie op verschillende individuen en organisaties tegelijkertijd. Mark en Henry identificeren vier mechanismen, te weten:

- 1 Agendabepaling (door afstemming met media, beleid en politiek);
- 2 Beleid georiënteerd leervermogen (de evaluatieresultaten gebruiken bij het vaststellen van beleid);
- 3 Beleid veranderen (beleid wordt aangepast als gevolg van de resultaten van de evaluatie);
- 4 Diffusie (veranderingen in beleid en programma's als gevolg van de resultaten van de evaluatie bij organisaties die buiten het onderzoeksgebied van de evaluatie vallen).

In de praktijk komt het vaak voor dat de invloed van evaluaties zich beweegt langs de drie verschillende niveaus (individueel, interpersoonlijk en collectief). Een individu kan op basis van de resultaten van een evaluatie veranderen van mening (eerste niveau). Deze mening kan met de nodige overtuigingskracht overgebracht worden op een collega (tweede niveau) waarna uiteindelijk het beleid van een organisatie kan worden bijgesteld (derde niveau). Doordat deze invloed zich kan voordoen op drie verschillende niveaus tegelijkertijd is het moeilijk om vast te stellen wat de invloed is geweest van een evaluatie op een bedrijf. Het raamwerk van Mark en Henry is echter interessant aangezien het aansluit op reeds bekende theorieën omtrent veranderprocessen en deze toepast op evaluaties. Door het combineren van deze veranderprocessen met de verschillende paden van evaluatie-invloeden kan in de toekomst gebruik

gemaakt worden van dit model om betere hypothesen te formuleren met betrekking tot de resultaten van evaluaties. Beoordelaars zullen, door gebruik te maken van dit model, in kaart moeten brengen welke personen door de evaluatie worden beïnvloed. Welke individuen met elkaar in gesprek zullen gaan na aanleiding van de resultaten van de evaluatie en welke organisatie dit raakt. Het proces om al deze personen en organisaties te identificeren kan gezien worden als een eerste stap om te bepalen hoeveel invloed een evaluatie heeft, of gaat hebben.

2.1.4 Mark en Henry Comprehensive Theory of Evaluation Influence

Mark en Henry (2004, p. 37) geven aan dat er bepaalde analogieën bestaan tussen de huidige theorieën omtrent benutting van evaluaties en de logische modellen met betrekking tot de gebruikte programma-theorieën, waarbij de programma's onderwerp zijn van evaluatie. In het verlengde hiervan hebben zij het conceptuele model van Cousins (2003, pp. 245-266) gebruikt, de evaluatietheorie ingepast en de genoemde analogie verder geëxpliciteerd. De benutting van evaluaties wordt dan beïnvloed door:

- Evaluatie-input (dit is inclusief de evaluatiecontext en de besluitvorming/beleidsetting);
- Evaluatieactiviteiten (datacollectie en analyse, stakeholder betrokkenheid, communicatie en disseminatie);
- Evaluatieoutput (kennis);
- Evaluatiebenutting (procesgebruik en/of resultaatgebruik).

Mark en Henry (2004, p. 37) leggen uit dat de huidige theorieën omtrent de benutting van evaluatie tegelijkertijd verarmd als overwoekerd zijn. Verarmd omdat de huidige modellen vooral omschrijven welke soorten benutting er zijn en hoe deze voorspeld kunnen worden. Ze missen een essentieel deel met betrekking tot welke onderliggende mechanismen in gang gezet worden om de resultaten van de evaluatie te benutten. Het identificeren van de onderliggende processen die mede de invloed van een evaluatie bepalen is niet alleen belangrijk voor de evaluatietheorie en haar onderzoeken maar ook voor de praktijk. Het begrijpen van deze processen zou er toe kunnen leiden dat er beter invloed uitgeoefend wordt en de evaluaties effectiever worden doorgevoerd. Met overwoekerd wordt bedoeld dat het relatief simpele concept benutting alternatieve, gedeeltelijk overlappende, maar wel tegenstrijdige betekenissen heeft gekregen. De verschillende vormen van benutting (instrumenteel, conceptueel, symbolisch en proces) overlappen elkaar en zijn niet strikt gescheiden. Een mooi voorbeeld hiervan is dat proces-benutting kan leiden tot het ondernemen van actie of het creëren van een bepaalde begripsvorming. Dit komt overeen met instrumentele en conceptuele benutting. Tevens is benutting een 'overwoekerd bos' omdat reeds vele verschillende posities binnen de scope van benutting zijn bepleit. Sommige onderzoekers beperken het benutten van een evaluatie tot de periode waarin de evaluatie plaatsvindt (Alkin, 2003). Anderen trekken het domein van evaluatiebenutting weer breder en nemen ook andere veranderingen mee. Uiteindelijk resulteert dit in een niet algemeen geaccepteerde definitie van benutting en wat het precies betekent wanneer een evaluatie wordt benut.

Een eerdere genoemde reactie is het gebruik maken van een nieuw concept, te weten: evaluatie-invloed (Kirkhart, 2000). Maar volgens Mark en Henry (2004, p. 39) is dit niet genoeg, namelijk: "it is essential to work to develop a more detailed and more specific framework and terminology of Influence as a way to clear out some of the conceptual thicket that has grown around the concept of use". In Kirkhart (Kirkhart,

2000) ontbrak een gedetailleerde beschrijving van de mechanismen die gebruikt kunnen worden door een evaluatie om invloed uit te oefenen. Mark en Henry (2003) hebben dit wel gedaan en classificeerden hun mechanismen in drie verschillende analyseniveaus, te weten individueel, interpersoonlijk en collectief, zie fig. 2. Vervolgens hebben zij in dit model de geïdentificeerde mechanismen geclassificeerd in vier verschillende typen, te weten: generieke invloedprocessen, cognitieve en affectieve processen, motiverende processen en gedragsprocessen.

Mark en Henry (2004) hebben een model ontwikkeld (zie tabel 1) waarin duidelijk de verschillende rollen van de onderliggende mechanismen worden benoemd en zichtbaar zijn. Door aandacht te schenken aan deze mechanismen en hun impact op de processen kunnen evaluaties in de toekomst beter gepland worden. Beoordelaars kunnen beter anticiperen op de processen die als gevolg van de evaluatie in gang worden gezet. Dit kunnen meerdere processen tegelijkertijd zijn waarbij processen ook getriggerd kunnen worden door andere processen (niet de evaluatie zelf). Elk individueel proces kan een korte termijn, intermediaire of lange termijn evaluatieresultaat zijn in de vaak complexe weg naar social betterment.

<i>Type of proces / outcome</i>	<i>Level of Analysis</i>		
	<i>Individual</i>	<i>Interpersonal</i>	<i>Collective</i>
General influence	Elaboration Heuristics Priming Skill acquisition	Justification Persuasion Change agent Minority-opinion influence	Ritualism Legislative hearings Coalition formation Drafting legislation Standard setting Policy consideration
Cognitive en affective	Salience Opinion/attitude valence	Local descriptive norms	Agenda setting Policy-oriented learning
Motivational	Personal goals and aspirations	Injunctive norms Social reward Exchange	Structural incentives Market forces
Behavioral	New skill performance Individual change in practise	Collaborative change in practise	Program continuation, cessation, or change Policy change Diffusion

Tabel 1. Alternatieve mechanismen welke evaluatie-invloed bemiddelen

Ook de meer klassieke vormen van evaluatiebenutting kunnen geplaatst worden in dit model. Instrumentele en conceptuele benutting zijn dan een type verandering die plaatsvindt als gevolg van een evaluatie. Hierbij kan instrumenteel uitgedrukt worden als verandering door actie en conceptueel als verandering door gedachte. Conceptuele benutting komt overeen met de cognitieve en affectieve processen. Instrumentele benutting komt overeen met de gedragsprocessen.

Symbolische benutting is niet zozeer te verbinden aan een resultaat of effect van een evaluatie maar eerder aan de intentie. Symbolische benutting daarentegen correspondeert met interpersoonlijke en collectieve processen van invloed. Symbolische benutting neemt in het model een relatief kleine plek in. Dit is niet verwonderlijk aangezien symbolische benutting weinig aandacht krijgt in de literatuur.

Procesbenutting wordt gedefinieerd door de bron van invloed: het proces. Bij procesbenutting heeft de verandering plaatsgevonden als gevolg van deelname aan het evaluatieproces in plaats van op basis van de evaluatieresultaten. Procesbenutting kan daarom per definitie niet vertaald worden in één of meerdere mechanismen. In de taal van logische modellen wordt de invloed van procesgebruik geactiveerd door evaluatieactiviteiten en niet de resultaten.

Dit model van Mark en Henry (2004, p. 41) elimineert de verwarring die kan ontstaan wanneer onderzoekers eenzelfde term gebruiken om te verwijzen naar verschillende veranderprocessen op de verschillende analyseniveaus. Vervolgens zijn Mark en Henry (2004, p. 45) gestopt met het verder categoriseren van de processen en zijn ze deze processen in gaan bedden in een nieuwe veelomvattende theorie van evaluatie-invloed. In deze theorie wordt het conceptuele model van Cousins (2003) gecombineerd met de geïdentificeerde verandermechanismen. Dit leidt uiteindelijk tot de volgende schematische weergave, zie figuur 3. Hierin is duidelijk te zien welke rol de verandermechanismen spelen. Deze veranderprocessen bemiddelen eigenlijk tussen de evaluatieactiviteiten en de uiteindelijke evaluatieresultaten. Waar in het verleden de veranderprocessen niet of nauwelijks in kaart zijn gebracht in relatie tot evaluaties wordt nu duidelijk welke beslissende rol zij kunnen spelen in het evaluatieproces.

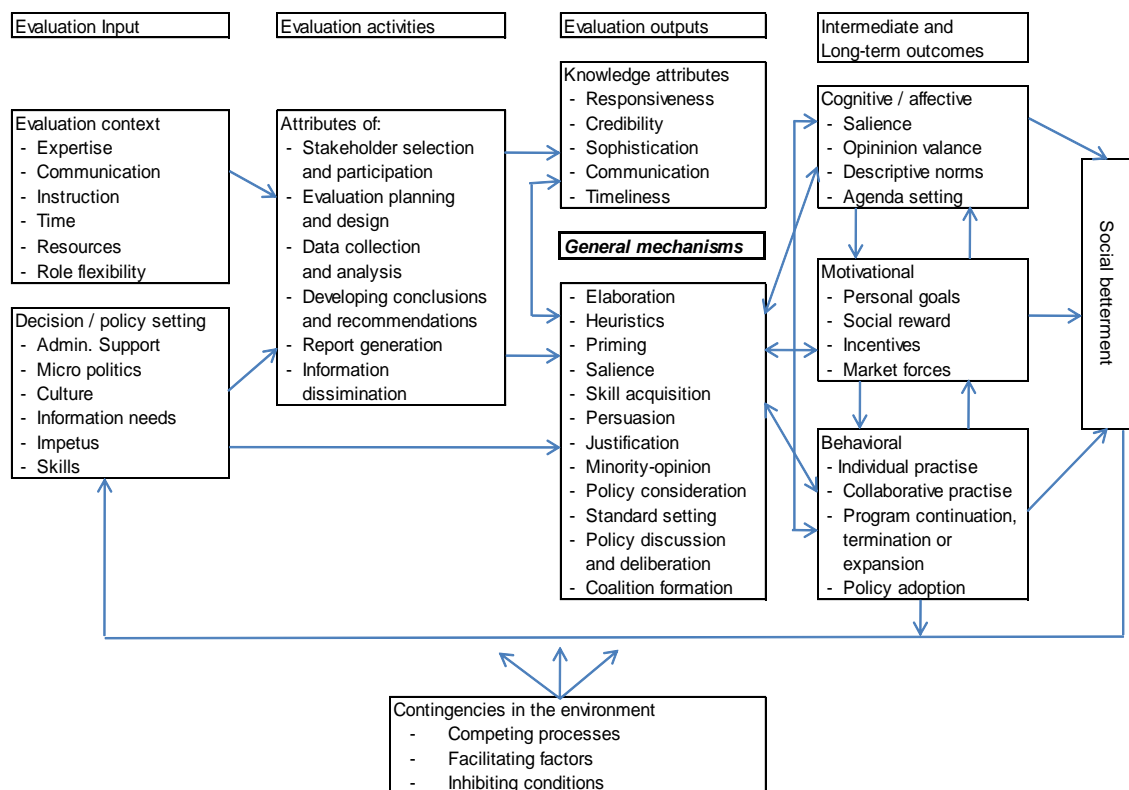


Fig. 3 Schematisch Theory of Evaluation Influence, Mark en Henry (2004)

Een voordeel van dit model is dat de paden van invloed visueel gemaakt worden. Er is een duidelijke bidirectionele link te zien tussen de drie gestelde niveaus van evaluatieresultaten en de onderliggende verandermechanismen. Dit geeft duidelijk weer dat de invloed van evaluaties zich kan bewegen langs de drie verschillende niveaus (Mark & Henry, 2003, p. 305). De paden naar evaluatie-invloed zijn altijd complex en bestaan uit meerdere verandermechanismen. In het verleden hebben onderzoeken naar het gebruik van evaluatie zich vaak beperkt tot enkele belangrijke sleutelfiguren, personen die beleidsbepalend zijn binnen organisaties. Deze gaven vaak aan dat evaluaties beperkte of minimale invloed hebben gehad op het gevoerde beleid. Melkers en Willoughby (2001, pp. 54-64) hebben met hun onderzoek echter aangetoond dat wanneer gedegen gekeken wordt naar de complexere paden van evaluatie-invloed de evaluaties meer invloed hebben gehad op het beleid dan de sleutelfiguren hadden verwacht of geconstateerd. Melkers en Willoughby (2001, pp. 54-64) hebben het belang van het onderkennen van deze paden aangetoond. Toekomstige evaluaties kunnen hierop inspelen waardoor hun invloed kan worden vergroot. Maar welke personen initiëren deze paden van invloed? Betreft het personen die actief betrokken zijn in het evaluatieproces of zijn het personen die alleen de resultaten van een evaluatie ontvangen?

2.2 Factoren die de benutting van evaluaties beïnvloeden

In deze paragraaf wordt het theoretische kader beschreven van de wijze waarop verschillende factoren van invloed kunnen zijn op de benutting van evaluaties. Hiermee wordt antwoord gegeven op de derde deelvraag van het onderzoek.

2.2.1 Overzicht van relevante onderzoeken en literatuur

In de jaren tachtig en negentig van het vorige decennium zijn er 5 veelomvattende literatuurstudies geweest naar de benutting van evaluaties (Alkin, 1985); (Beyer & Trice, 1982); (Cousins & Leithwood, 1986); (King & Thompson, 1983); en (Leviton & Hughes, 1981). De meest recente uit die periode is die van Shulha en Cousins (1997). Om alle factoren die bewezen invloed hebben op de benutting van evaluaties te kunnen categoriseren, zijn onderzoekers begonnen met deze factoren te groeperen in verschillende categorieën. Alkin (1985, p. 12) identificeert drie categorieën van factoren die de benutting van evaluaties kunnen beïnvloeden, te weten:

1. Menselijke factoren, zoals de kenmerken van de beoordelaar en de gebruiker, hun interesse in de evaluatie en hun professionele ervaring en positie binnen de organisatie;
2. Contextuele factoren, zoals de setting van het programma, financiële omstandigheden en het sociale en politieke klimaat;
3. Evaluatiefactoren die refereren naar de evaluatie zelf. De informatie die de evaluatie bevat en de manier waarop deze informatie is vergaard.

De eerste onderzoeken naar factoren beperkten zich meer tot het benoemen en omschrijven van de geïdentificeerde factoren. Leviton en Hughes (1981, pp. 525-548) hebben zich tijdens hun literatuuronderzoek naar de benutting van evaluaties geconcentreerd op de periode 1965 tot 1980 en hebben op basis van de beschikbare literatuur vijf categorieën geïdentificeerd die van invloed zijn op de benutting van evaluatie, te weten:

1. Relevantie van de gekozen deelnemers, de te onderzoeken resultaten en de rapportageperiode;
2. Effectiviteit van de communicatie tussen de beoordelaars en potentiële gebruikers;
3. Vaardigheid om de resultaten van de evaluatie effectief te kunnen verwerken;
4. Geloofwaardigheid van de beoordelaar;
5. Gebruikersmedewerking en rol als pleitbezorger/verdediger van de evaluatie. Dit refereert naar de betrokkenheid van gebruikers om de resultaten van de evaluatie te gebruiken in het besluitvormingsproces.

Cousins en Leithwood (1986, pp. 331-364) hebben zich tijdens hun literatuuronderzoek naar benutting van evaluaties geconcentreerd op de periode 1971 tot en met 1986. Deze periode van onderzoek is grotendeels gelijk aan die van Leviton en Hughes. Veel van de factoren die zij hebben geïdentificeerd komen dan ook overeen met die van Leviton en Hughes. Wel hebben zij gekozen voor een andere manier van presenteren. Zij zijn de eersten geweest die een conceptueel model hebben ontwikkeld dat verder reikt dan een simpele omschrijving van factoren. Zij identificeren 12 factoren die van invloed zijn op de benutting van een evaluatie en delen deze in twee verschillende categorieën in, namelijk de evaluatiesetting en de besluitvorming- en beleidsetting van het evaluatieprogramma.

Groot belang mag worden gehecht aan hun methode voor het toekennen van waardes aan de verschillende factoren. Door het gebruik van deze waardes en dus de volgordelijkheid van belang kregen de beoordelaars een indicatie van de hoeveelheid tijd welk zij zouden moeten besteden aan de verschillende factoren. Zij hebben een methode ontwikkeld waarmee een relatie-index is gecreëerd die de sterkte van elke relatie tussen een onafhankelijke en afhankelijke variabele vergelijkt. Deze index had een driedelig doel. Ten eerste maakt het een inschatting van alle factoren over alle soorten benutting en potentie voor benutting. Vervolgens worden de verschillen in sterkte van invloed beoordeeld tussen de factoren. Om als laatste te kunnen constateren welke factoren de meeste invloed uitoefenen. Zij vonden dat benutting het meest vanzelfsprekend was wanneer:

- a) Evaluaties gebruik maken van de juiste methode en complexiteit;
- b) Suggesties toepasselijk en van belang zijn voor participanten;
- c) De uitkomsten overeenkomen met de gebruikersverwachtingen en geloof;
- d) Gebruikers direct betrokken zijn bij de totstandkoming van de evaluatie en de resultaten onderschrijven;
- e) Het data rapport relevant is voor het probleem van de gebruiker.

Shulha en Cousins (1997) hebben zich tijdens hun literatuuronderzoek naar benutting van evaluaties geconcentreerd op de periode tussen 1986 en 1997 en hebben op basis van de beschikbare literatuur 5 belangrijke ontwikkelingen in de theorie, research en praktijk gevonden. De eerste heeft betrekking op het verder benadrukken van de context van een evaluatie. De context is direct gelieerd aan het kunnen begrijpen en verklaren van benutting. Beoordelaars worden aangemoedigd om meer te leren van de cultuur en structuur van de organisaties die participeren in de evaluatie. De tweede ontwikkeling is het toevoegen van procesbenutting als een vorm van benutting. De derde ontwikkeling is de focus op de verschillen tussen individuen en organisaties en hoe zij evaluatie benutten. Verdere begripsvorming omtrent misbruik van evaluatieresultaten is de vierde ontwikkeling en de laatste is de veranderende rol van de beoordelaar. Deze rol heeft additionele verantwoordelijkheden gekregen. Hierbij moet men denken

aan faciliteren, plannen en opleiden. Elk van deze ontwikkelingen zijn momenteel populaire onderwerpen in de literatuur (Amo & Cousins, 2007); (Fetterman, 2003); (Mark & Henry, 2003); (Patton M. , 2007).

De meest recente literatuurstudie naar de benutting van evaluaties is die van Johnson et Al. (Johnson, Greenesid, Toal, King, Lawrenz, & Volkov, 2009). Als basis voor hun review, de onderliggende structuur, werd gebruik gemaakt van het model van Shulha en Cousins (1997) en het model van Cousins en Leithwood (1986). Johnson et al. (2009) hebben het gebruik van andere modellen (bijvoorbeeld op basis van invloed) overwogen, maar nader onderzoek heeft uitgewezen dat de bestaande modellen nog steeds het meest veelomvattend, beste gedefinieerd en concreet zijn. Wel hebben Johnson et Al. (2009) geconstateerd dat de literatuurstudies van Shulha en Cousins (1997) en Cousins en Leithwood (1986) niet volledig waren. Niet alle bestaande onderzoeken waren opgenomen en ook het aspect 'kwaliteit van het onderzoek' was niet opgenomen.

Zo kan het zijn dat onderzoeken waarbij de gekozen methodologie dubieus is, eenzelfde waarde wordt toegekend als onderzoeken die wel van hoogwaardige academische kwaliteit zijn. Om niet dezelfde fout te maken als hun voorgangers hebben Johnson et Al. (2009) alle empirische onderzoeken naar de benutting van evaluaties meegenomen (ook de proefschriften) en hebben op basis van voorgedefinieerde criteria de kwaliteit van de methodologie beoordeeld. Hun onderzoek leverde in eerste instantie ruim 600 artikelen en 48 proefschriften op. Na een eerste high level screening bleven er nog 321 over die voldeden aan de voorgedefinieerde criteria van Johnson et al. (2009). Na het doornemen van de samenvattingen identificeren Johnson et al. (2009) 98 publicaties die in detail doorgenomen moesten worden. Op basis van een volledige review van de tekst werden deze publicaties nogmaals getoetst aan de voorgedefinieerde criteria. Uiteindelijk bleven er na dit proces nog 47 onderzoeken over. Deze onderzoeken kregen een kwaliteitsindicatie en 41 van de 47 onderzoeken scoren een voldoende. Deze overgebleven 41 onderzoeken vormden de basis van de literatuurstudie van Johnson et Al (2009). De belangrijkste bijdrage van het onderzoek van Johnson et Al. (2009) is het toevoegen van de bekwaamheid van de evaluator aan de factoren behorende tot de evaluatiesetting en het toevoegen van een nieuwe categorie 'stakeholder betrokkenheid'.

Uiteindelijk hebben Johnson et Al. (2009) een model opgeleverd waarin alle 41 onderzoeken worden toegewezen aan specifieke categorieën. Tevens is hierbij aangegeven wat de focus is van het onderzoek, het aantal analyse-eenheden en de categorie van benutting. In de toekomst kunnen onderzoekers gebruik maken van deze informatie om vergelijkbare evaluatieonderzoeken te vinden en mogelijk verder in te zoomen op geconstateerde verschillen. Johnson et Al. (2009) hebben in de categorie 'factoren die van invloed zijn op de benutting van evaluaties die behoren tot evaluatiesetting en besluitvorming- en beleidsetting', ook het onderzoek van Bober en Bartlett meegenomen (2004).

2.2.2 Onderzoek Bober & Bartlett (2004)

In 2004 hebben Bober en Bartlett een onderzoek gedaan naar de benutting van evaluaties van training-programma's bij een viertal corporatieve universiteiten. Een corporatieve universiteit is een separate entiteit binnen een onderneming die verantwoordelijk is voor het ontwikkelen en implementeren van trainingprogramma's voor medewerkers. Doelstelling van dit onderzoek was tweeledig. Ten eerste het identificeren van medewerkers die de data van de evaluatie benutten en de reden hiervoor. Ten tweede

het identificeren van factoren die volgens deze medewerkers de benutting van de evaluatie beïnvloeden binnen deze corporatieve universiteiten.

Bober en Bartlett (2004) hebben als basis voor hun onderzoek het conceptuele model van Cousins en Leithwood (1986) gebruikt. Ze kozen dit model vanwege de volledigheid en toepasselijkheid voor het onderzoeken van factoren en andere aspecten met betrekking tot de benutting van evaluaties. Een additioneel voordeel van dit bestaande model is volgens Bober en Bartlett “the detailed way in which it was developed through a meta-analysis of existing frameworks to produce a model that built on (and synthesized) the most important components influencing utilization” (2004, p. 368). Dit model benoemt niet alleen de verschillende factoren maar voorziet ook in een manier om, op basis van feitelijke invloed, deze in een rangorde te plaatsen.

<i>Evaluatiesetting</i>	<i>Besluitvorming- en beleidsetting</i>
1. De kwaliteit van de evaluatie	1. De informatienoodzaak
2. De geloofwaardigheid van de evaluatie	2. De besluitvormingskenmerken van de organisatie
3. De relevantie van de evaluatie	3. Het politieke klimaat
4. De communicatie kwaliteit van de evaluatie	4. Concurrerende informatie
5. De resultaten van de evaluatie	5. Persoonlijke kenmerken van de gebruiker
6. De timing van de evaluatie, op tijd, te laat	6. Commitment naar de evaluatie toe

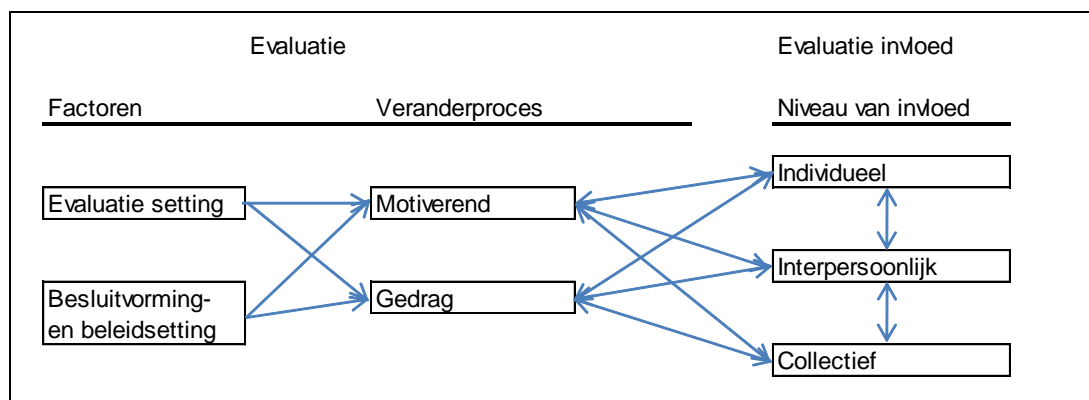
Uit het onderzoek van Bober en Bartlett (2004) bleek dat het model van Cousins en Leithwood toepasbaar was en dat door belanghebbenden 7 factoren werden geïdentificeerd die overeen kwamen met de factoren vanuit het Cousins en Leithwood model. Hierbij konden 6 van de 7 factoren gecategoriseerd worden onder de subcategorie ‘evaluatiesetting’.

2.3 Conceptueel model

Het conceptueel model dat op basis van het theoretisch kader is ontwikkeld is gebaseerd op de uitgangspunten van het model van Cousins en Leithwood (1986) en Mark en Henry (2004), zie figuur 2.1. In het model van Cousins en Leithwood (1986, pp. 347-348) zijn 12 verschillende factoren samengevat in 2 categorieën die van invloed zijn op de benutting van evaluaties. Deze 2 categorieën zijn de evaluatiesetting en de besluitvorming- en beleidsetting. De categorie evaluatiesetting kan verder onderverdeeld worden naar een zestal factoren, namelijk: kwaliteit, geloofwaardigheid, relevantie, communicatie, resultaten en timing van de evaluatie. De categorie besluitvorming- en beleidsetting kan ook verder onderverdeeld worden naar een zestal factoren, namelijk: informatienoodzaak, besluitvormingskenmerken, politiek klimaat, concurrerende informatie, persoonlijke kenmerken en commitment naar de evaluatie toe. De twee hoofdcategorieën factoren uit het model van Cousins en Leithwood vormen het eerste deel van het conceptuele model dat in dit onderzoek gebruikt wordt. Vervolgens zijn hier de drie verschillende niveaus van evaluatie-invloed van Mark en Henry (2004, pp. 40-41) aan toegevoegd, namelijk: individueel, interpersoonlijk en collectief.

Aan deze verschillende niveaus van evaluatie-invloed gaan volgens Mark en Henry (2004) de verandermechanismen vooraf. Juist deze verandermechanismen stelt de evaluatie in staat om invloed uit te oefenen. De invloed beweegt zich tussen deze niveaus en dit kan dan ook gezien worden als een interactief proces. Elk van de niveaus correspondeert met een veranderproces dat is bestudeerd in de sociale- en gedragswetenschappen. De geïdentificeerde veranderprocessen zijn: generieke invloed, cognitief en affectief, motiverend en gedragsmatig. In dit onderzoek zijn alleen de veranderprocessen 'motiverend' en 'gedragsmatig' relevant. In tegenstelling tot de andere twee veranderprocessen vereisen deze twee veranderprocessen een trigger, een actie/gebeurtenis, die voorafgaat aan het veranderproces. In dit onderzoek is de evaluatie deze trigger. Bij de generieke en cognitieve/affectieve veranderprocessen is het verband met de evaluatie minder duidelijk en kunnen andere factoren de oorzaak zijn van de veranderprocessen. Daarbij kan ook gesteld worden dat de motiverende en gedragsmatige veranderprocessen focussen op de korte tot middellange termijn en kennen dus, in relatie tot de evaluatie, een grotere mate van nauwkeurigheid dan de generieke en cognitieve/affectieve veranderprocessen. Deze laatste twee focussen met name op de middellange tot lange termijn.

Dit onderzoek stelt ten eerste vast welke factoren van belang zijn voor het genereren van evaluatie-invloed. Hierbij worden de factoren toegewezen aan de categorie evaluatiesetting of de categorie besluitvorming- en beleidsetting. Vervolgens is onderzocht en in kaart gebracht wat de feitelijke evaluatie-invloed is bij WML en Brabant Water en hoe deze evaluatie-invloed tot stand is gekomen. Welke verandermechanismen hebben aan de basis gestaan van deze evaluatie-invloed? Uiteindelijk wordt de feitelijke evaluatie-invloed in verband gebracht met de aan- of afwezigheid van bepalende factoren. Een beter begrip van de factoren die van belang zijn voor evaluatie-invloed kunnen in de toekomst leiden tot betere evaluaties die gericht invloed kunnen uitoefenen. Vooral de term *gericht* is van belang. Meer is niet altijd beter, een praktisch voorbeeld is communicatie. Meer communiceren kan in het geval van slechte communicatie leiden tot grotere onduidelijkheden. Gericht en beter communiceren is dan een veel betere oplossing.



Figuur 2.1 Conceptueel model op basis van theoretisch kader

3 METHODOLOGIE

In dit hoofdstuk wordt de gekozen onderzoeksmethode, in relatie tot de probleemstelling van dit onderzoek, onderbouwd. Aan bod komen de keuze en kwaliteit van de meetinstrumenten en de wijze waarop de variabelen en concepten zijn geoperationaliseerd en gemeten in het onderzoek.

3.1 Keuze van de onderzoeksstrategie

Volgens Yin (2009, p. 8) geven de antwoorden op de volgende 3 vragen voldoende inzicht voor het bepalen van de meest geschikte onderzoeksstrategie, de opzet van de onderzoeksmethoden en de uitvoering van het onderzoek (inclusief dataverzameling en dataverwerking):

1. Wat is de vorm van de onderzoeksvraag?
2. Vereist het onderzoek een bepaalde mate van gedragscontrole?
3. Ligt de focus van het onderzoek op de tegenwoordige tijd?

In dit onderzoek is vooral gekeken naar het hoe en waarom: er is een verband gelegd tussen betrokkenheid en evaluatie-involed. Het is voor het onderzoek niet noodzakelijk om controle te hebben over het gedrag van participanten en de centrale onderzoeksvraag concentreert zich op de tegenwoordige tijd. De gevalsstudie is derhalve de meest voor de handliggende onderzoeksmethode voor dit onderzoek. Dit is in overeenstemming met de methode van verklarend onderzoek, het sluit aan bij het hoe en waarom van een gevalsstudie in plaats van het hoe, wat, waar en hoeveel van een survey.

Bij het ontwerpen van een gevalstudie is het volgens Yin noodzakelijk om vooraf een onderscheid te maken tussen een enkelvoudige of een meervoudige gevalsstudie (Yin, 2009, pp. 46-60). Yin (2009, pp. 47-50) schrijft een vijftal kenmerken toe aan een enkelvoudige gevalstudie. Hierbij wordt rekening gehouden met het feit dat een enkelvoudige gevalsstudie analoog is aan een enkelvoudig experiment en dat veel van de condities die gelden bij enkelvoudige experimenten ook van toepassing zijn op enkelvoudige gevalsstudies. De vijf redenen om een gevalsstudie als enkelvoudig te benoemen zijn:

1. Wanneer het een kritische casus representeert voor het testen van een goed geformuleerde theorie. Een casus die voldoet aan al deze eisen betreffende het testen van een theorie, kan de theorie bevestigingen, ontkrachten of zelfs uitbreiden;
2. Wanneer het een extreme of unieke casus representeert;
3. Wanneer het een typische of representatieve casus representeert waarbij het doel is om de condities en omstandigheden vast te kunnen leggen die leiden tot alledaagse situaties;
4. Wanneer het een openbarende casus representeert. Door omstandigheden krijgt een onderzoeker de mogelijkheid om een situatie te observeren en analyseren die voorheen nooit toegankelijk of beschikbaar is geweest;
5. Wanneer het een longitudinale casus representeert waarbij een zelfde casus op twee verschillende momenten in de tijd wordt bestudeerd.

De eerst genoemde reden is van toepassing op dit onderzoek en derhalve is gekozen voor een enkelvoudige gevalsstudie (Yin, 2009). Bober & Bartlett (2004) hebben reeds aangetoond in hun

onderzoek dat de evaluatiesettingfactoren en de besluitvorming- en beleidsettingfactoren van invloed zijn op de benutting van evaluaties. Dit onderzoek heeft als doelstelling om deze theorie uit te breiden. Tevens is de derde genoemde reden van toepassing op dit onderzoek. Het gaat namelijk om een typische representatieve casus die leidt tot alledaagse situaties. Vervolgens geeft Yin (2009, p. 50) aan dat bij een enkelvoudige casus de keuze gemaakt moet worden tussen een holistische of ingebedde enkelvoudige gevalsstudie.

3.1.1 Holistisch of ingebedde enkelvoudige gevalsstudie?

Een enkelvoudige gevalsstudie kan bestaan uit meerdere analyse-eenheden. Dit gebeurt wanneer binnen de casus aandacht wordt besteed aan een subeenheid of verschillende eenheden (Yin, 2009, p. 50). Deze ingebedde eenheden kunnen op basis van een steekproef- of clustermethode geselecteerd worden. Er wordt dan gesproken van een ingebedde enkelvoudige gevalsstudie. Het tegenovergestelde gebeurt wanneer de gevalsstudie het mondiale karakter van een organisatie of programma onderzoekt. Er is dan sprake van een holistische enkelvoudige gevalsstudie. In dit onderzoek betreft het duidelijk een ingebedde enkelvoudige gevalsstudie.

Een groot voordeel van de meer complexe ingebedde methode is dat het hebben van meerdere subeenheden veel mogelijkheden creëert voor extensieve analyse. Hierdoor kunnen de inzichten in de enkelvoudige gevalsstudie worden verscherpt. Daartegenover staat het nadeel dat wanneer er te veel focus wordt gelegd op de verschillende subeenheden, het holistische aspect van de gevalsstudie kan worden verwaarloosd. De gevalsstudie verandert dan van oriëntatie. Indien deze verandering gewenst is moet dit uitdrukkelijk worden benoemd en dient de relatie met het originele onderzoek te worden gelegd.

3.1.2 Keuze van de gevalsstudie

Om aan te kunnen tonen dat de evaluatiesettingfactoren en de besluitvorming- en beleidsettingfactoren van belang zijn voor de evaluatie-invloed, is het van belang dat gekozen wordt voor een evaluatie waaraan verschillende gelijksoortige organisaties participeren. Als casus is gekozen voor de benchmark 'Water in Zicht' binnen de Nederlandse drinkwatersector.

In de benchmark worden de prestaties van alle drinkwaterbedrijven op procesniveau met elkaar vergeleken. Vanwege het belang dat de drinkwaterbedrijven hechten aan de benchmark hebben zij hier altijd aan deelgenomen, ook voordat deelname verplicht werd door de Nederlandse Overheid.

Kirkhart (2000) gaf al aan dat invloed moeilijk te meten is en op verschillende momenten in de tijd kan worden geconstateerd. Deelname aan een evaluatieproces en de evaluatieresultaten zullen dus zelden direct en traceerbaar leiden tot zichtbare veranderingen binnen een onderneming. Waarbij het tevens mogelijk is dat deze veranderingen niet zichtbaar zijn maar onderhuids spelen en betrekking hebben op gedrag en motivatie van medewerkers (Mark & Henry, 2004). Dit zijn de eerder toegelichte verander-mechanismen (zie hoofdstuk 2).

De benchmark binnen de drinkwatersector is een jaarlijks terugkerend proces met een lange doorlooptijd. Evaluatie-invloed kan hierdoor moeilijk toegewezen worden aan één specifieke benchmark. De evaluatie-invloed kan zich hebben ontwikkeld over verschillende jaren. De keuze van het moment

waarop het onderzoek plaatsvindt kent hierdoor geen beperkingen. Dit betekent dat de invloed van de benchmark over meerdere jaren is onderzocht.

Er is echter wel gekozen voor een specifiek meetmoment, namelijk vlak na de presentatie van de benchmarkresultaten. Op dat moment wordt aan de directie van de deelnemende bedrijven verantwoording afgelegd omtrent de behaalde benchmarkresultaten van het afgelopen jaar en is de aandacht voor de benchmark binnen de organisaties het grootst. Tevens is dan ook duidelijk welke mogelijke vervolgacties voortvloeien uit deze resultaten en welke medewerkers betrokken worden in het vervolgtraject.

3.1.3 Keuze van de analyse-eenheden

Een enkelvoudige casus kan bestaan uit meerdere analyse-eenheden. Dat is zeker het geval voor de casus in dit onderzoek. Om te kunnen aantonen dat de evaluatiesettingfactoren en de besluitvorming- en beleidsettingfactoren van invloed zijn op de invloed van een evaluatie zijn meerdere analyse-eenheden noodzakelijk. De analyse-eenheden kunnen niet random gekozen worden, maar moeten aan een tweetal eisen voldoen. Ten eerste moeten de bedrijven deelnemen aan de benchmark. Ten tweede dienen de geselecteerde bedrijven in grote lijnen vergelijkbaar te zijn. Op die manier worden mogelijke andere factoren van invloed - die te maken hebben met het type bedrijf - uitgesloten.

Op basis van de genoemde argumenten zijn als analyse-eenheden WML (Watermaatschappij Limburg) en Brabant Water gekozen. Beide nemen deel aan de benchmark. Daarnaast zijn beide bedrijven in grote lijnen vergelijkbaar, het zijn drinkwaterbedrijven die voornamelijk grondwater gebruiken.

3.1.4 Keuze van de waarnemingseenheden

Een analyse-eenheid kan bestaan uit meerdere waarnemingseenheden. Dat is zeker het geval voor de analyse-eenheden in dit onderzoek. Om te kunnen aantonen dat de evaluatiesettingfactoren en de besluitvorming- en beleidsettingfactoren van invloed zijn op de invloed van een evaluatie zijn meerdere waarnemingseenheden noodzakelijk. Alle waarnemingseenheden moeten een rol hebben binnen het benchmarkproces of de resultaten van de benchmark ontvangen. Vervolgens dienen er bepaalde overeenkomsten te zijn tussen de geselecteerde waarnemingseenheden van beide analyse-eenheden. Het komt de validiteit van het onderzoek niet ten goede wanneer bij de eerste analyse-eenheid alle waarnemingseenheden uit staffunctionarissen bestaan en bij de tweede alleen maar uit operationele managers. Binnen dit onderzoek bestaan de waarnemingseenheden dan ook zo veel mogelijk uit gelijksoortige functionarissen.

3.2 Dataverzameling

Door gebruik te maken van meerdere bronnen of waarnemingsmethodieken kunnen er meerdere informatielijnen ontstaan die tot eenzelfde conclusies komen. Dit kan ingezet worden om de verkregen resultaten te verifiëren. Dit heet triangulatie. Triangulatie wordt vaak gebruikt om te impliceren dat er meer dan twee dataverzamelingsmethode zijn gebruikt in een onderzoek om een bepaald resultaat te controleren. In de praktijk zal de betrouwbaarheid van een onderzoek groter zijn, als meerdere

onderzoeksmethodes, door middel van triangulatie het zelfde resultaat geven. Een goede gevalsstudie zal dan ook meerdere bronnen tegelijkertijd raadplegen.

In zijn boek beschrijft Yin (2009, pp. 101-102) een zestal veel voorkomende - en complementaire - bronnen om data te verzamelen, namelijk:

1. Documentstudie;
2. Gearchiveerde bestanden (papier als digitaal);
3. Interviews;
4. Directe observatie (passieve rol);
5. Participatie observatie (actieve rol in gevalsstudie);
6. Fysiek artefact.

Dit onderzoek maakt gebruik van documentstudie en interviews. De benchmark is de afgelopen tien jaar elk jaar uitgevoerd. Het proces, de planning, de betrokken personen en hun verantwoordelijkheden zijn dan ook goed gedocumenteerd. De factoren die volgens de belanghebbenden van belang zijn voor de evaluatie-invloed zijn echter persoonlijk en kunnen alleen door middel van interviews achterhaald worden. Het voeren van interviews is de meest geschikte methode om de evaluatie-invloed en onderliggende verandermechanismen in kaart te brengen.

Toelichting opzet, werkwijze en verwerking data uit interviews

Om de factoren en oorzaken van invloed te achterhalen is gebruik gemaakt van open vragen. Het interview is semigestructureerd opgezet en uitgevoerd. Dat wil zeggen dat de vragen vooraf zijn bepaald en vastgesteld, waarbij afhankelijk van de antwoorden ook verdiepende vragen in de routing van de vragenlijst zijn opgenomen. Aan deze interviews ligt een vragenlijst vanuit het theoretische kader ten grondslag. Deze lijst is niet bedoeld om de geïnterviewde te leiden naar bepaalde antwoorden, maar diende vooral als rode draad om te bewaken dat gedurende het interview alle aspecten aan bod kwamen. Er zijn in totaal dertien belanghebbenden geïnterviewd; zeven respondenten bij Brabant Water en zes respondenten bij WML.

Zes van de dertien belanghebbenden zijn afdelingshoofd (twee regiomanagers bij WML en een afdelingshoofd bedrijfsvoering productie, distributie, klantenservice en financieel economische zaken bij Brabant Water) en één belanghebbende is dit lang geweest (voormalig afdelingshoofd corporate control bij WML). Deze personen zijn allen verantwoordelijk voor een of meerdere processen die benoemd en gerapporteerd worden in de benchmark. Twee belanghebbenden zijn directeur (Brabant Water en WML) en twee belanghebbenden zijn sectormanager (bij Brabant Water manager Productie en bij WML manager Klant en Operaties). Zij zijn vanuit hun functie binnen de betreffende organisaties verantwoordelijk voor de resultaten van de benchmark. De sectormanagers zijn verantwoordelijk voor specifieke processen van de benchmark, de directeuren voor de totale benchmark. Een van de belanghebbenden is lid van de Raad van Commissarissen bij WML. De resultaten van de benchmark worden jaarlijks gerapporteerd aan de Raad van Commissarissen. De laatste belanghebbende heeft een meer uitvoerende functie bij Brabant Water (consulent financieel economische zaken) maar heeft een belangrijke rol bij het tot stand komen van de benchmark (input).

Tijdens het interview is op basis van de gegeven antwoorden indien nodig verder ingezoomd op het hoe en waarom. Met name praktijkvoorbeelden zijn in dit onderzoek van belang omdat evaluatie-invloed tastbaar moet zijn en aangetoond kan worden aan de hand van praktijkvoorbeelden. Op die manier kan onderzocht worden hoe deze tot stand is gekomen en of deze invloed ook door documentstudie bevestigd wordt. Om de betrouwbaarheid van de data te vergroten zijn de antwoorden van de verschillende belanghebbenden binnen dezelfde organisatie met elkaar vergeleken. Daar waar nodig zijn achteraf extra toelichtingen gevraagd aan de geïnterviewden. De resultaten van de gesprekken zijn samengevat en geverifieerd bij de geïnterviewde (respondent). In de digitale versie van dit onderzoek zijn de interviewverslagen wegens geheimhouding niet in de bijlagen opgenomen. De interviewverslagen zijn wel ter beschikking gesteld aan de faculteit managementwetenschappen van de OU.

In onderliggende subparagrafen is per deelvraag beschreven welke methoden zijn gebruikt om data te verzamelen en hoe triangulatie heeft plaatsgevonden. Wanneer er bij onderstaande subparagrafen wordt gesproken over interviews wordt tevens verwezen naar bovenstaande toelichting, waarin de opzet, uitvoering en verwerking van data is beschreven.

3.2.1 De methoden bij deelvraag 1, 2, 3 en 4

De eerste vier deelvragen zijn beantwoord middels het uitvoeren van een literatuurstudie. Op basis van de uitkomsten van de literatuurstudie is het theoretische kader van dit onderzoek geformuleerd (zie hoofdstuk 2). Dit theoretische kader dient als uitgangspunt voor de andere deelvragen.

3.2.2 De methoden bij deelvraag 5 en 6

Deelvraag 5 beperkt zich tot de factoren die behoren tot de evaluatiesettingfactoren. Deelvraag zes beperkt zich tot de factoren die behoren tot de besluitvorming- en beleidsettingfactoren. Bij beide vragen gaat het om de mening, ervaring en beleving van belanghebbenden. In dit geval is het interview dan ook de meest geschikte (data)bron.

3.2.3 De methoden bij deelvraag 7 en 8

Deelvraag zeven gaat over het wel of niet aanwezig zijn van de verschillende niveaus van evaluatie-invloed. Om de verschillende niveaus van invloed te kunnen achterhalen is ten eerste in kaart gebracht of er evaluatie-invloed is. Dit kon alleen achterhaald worden op basis van interviews. Op basis van de antwoorden is vervolgens bepaald op welk niveau deze evaluatie-invloed heeft plaatsgevonden. Hoewel triangulatie in dit geval erg moeilijk was, is getracht om via de beschikbare documentatie de bewijzen van de genoemde evaluatie-invloed terug te vinden. Deelvraag acht gaat over de verandermechanismen die aan de basis van de evaluatie-invloed hebben gestaan. Het antwoord op deze deelvraag is met behulp van interviews achterhaald. Triangulatie was hier niet mogelijk.

3.2.4 De methoden bij deelvraag 9, 10 en 11

Bij de laatste drie deelvragen zijn verbanden gelegd tussen de verschillende aspecten uit het conceptueel model. Hiertoe zijn de resultaten van de deelvragen vijf tot en met acht geanalyseerd op basis van de theoretische beweringen die ten grondslag liggen aan het onderzoek (Yin, 2009). Met

behulp van de uitkomsten van de laatste drie deelvragen is de centrale vraag van dit onderzoek beantwoord.

3.3 Beperkingen van de onderzoeksresultaten

Achtereenvolgens wordt een opsomming gegeven van beperkende aspecten met betrekking tot de verkregen onderzoeksresultaten. Deze aspecten dienen meegenomen te worden bij de interpretatie van de resultaten in relatie tot de reikwijdte en omvang van het onderzoek, conform de vastgestelde aanpak en uitvoering.

Een groot aantal onderzoeken heeft zich de afgelopen tientallen jaren gefocust op factoren die de benutting van evaluaties beïnvloeden (Alkin, 1985); (Shulha & Cousins, 1997); (Leviton & Hughes, 1981). Dit onderzoek beperkt zich niet tot specifieke factoren maar vraagt in navolging van Bober & Bartlett (2004) de belanghebbenden welke factoren volgens hen belangrijk zijn om deze vervolgens via de methodiek van Cousins & Leithwood (1986) te classificeren en rangschikken. Dit onderzoek toont aan dat deze factoren niet alleen van belang zijn voor de benutting van evaluaties bij corporatieve universiteiten maar ook van betekenis zijn voor de evaluatie-invloed binnen de Nederlandse drinkwatersector. Andere factoren, die niet geclassificeerd kunnen worden volgens het model van Cousins en Leithwood (1986), maar wel de evaluatie-invloed beïnvloeden, worden wel benoemd maar zijn niet verder onderzocht. Naast de in dit onderzoek genoemde factoren kan de evaluatie-invloed dus ook door andere factoren worden beïnvloed.

In dit onderzoek is de keuze gemaakt om de invloed van evaluatie te beperken tot de drie niveaus uit het model van Henry en Mark: individueel, interpersoonlijk en collectief (2004, p. 41). Tevens worden in dit onderzoek alleen de motiverende en gedragsmatige veranderprocessen belicht. De generieke en cognitieve affectieve veranderprocessen worden niet meegenomen. De keuze voor het model van Mark & Henry (2004, p. 41) is gebaseerd op de volgende overwegingen. Dit model bouwt voort op eerdere kennis en literatuur omtrent evaluatie-invloed (Kirkhart, 2000, pp. 5-23); (Mark & Henry, 2003, pp. 293-314) waarbij ook de klassieke vormen van evaluatiebenutting geïmplementeerd kunnen worden.

In dit onderzoek is bewust gekozen voor evaluatie-invloed in plaats van de benutting van evaluaties. Door de wijze waarop de benchmark is vormgegeven binnen de te onderzoeken bedrijven is het onmogelijk om een duidelijk onderscheid te kunnen maken tussen proces- en resultaatbenutting. De benchmark is een jaarlijks terugkerende evaluatie met een doorlooptijd van bijna een jaar. Belanghebbenden kunnen hierdoor moeilijk onderscheid maken tussen proces- en resultaatbenutting. Het komt regelmatig voor dat de resultaten van de laatste benchmark reeds worden gepubliceerd terwijl de acties als gevolg van de vorige benchmark nog niet zijn afgerond.

De benchmark binnen de Nederlandse drinkwatersector is een jaarlijks terugkerend proces met een lange doorlooptijd. Evaluatie-invloed kan daardoor moeilijk toegewezen worden aan één specifieke Benchmark, maar kan zich ontwikkeld hebben over verschillende jaren. Toch is om de betrouwbaarheid van de bevindingen van dit onderzoek te vergroten, gekozen om de evaluatie-invloed te onderzoeken

één tot drie maanden na de presentatie van de benchmarkresultaten. Op dat moment wordt aan de directie van de deelnemende bedrijven verantwoording afgelegd omtrent de behaalde resultaten en is het belang van de benchmark binnen de deelnemende organisaties het grootst.

De drinkwaterbedrijven Brabant Water en WML zijn niet de enige bedrijven die participeren in de benchmark. Alle Nederlandse drinkwaterbedrijven nemen deel aan de benchmark. De onderzoeksresultaten beperken zich echter tot de eerder genoemde twee organisaties.

Om de betrouwbaarheid van dit onderzoek te vergroten zijn alleen belanghebbenden meegenomen die de volledige resultaten van de evaluatie ontvangen. Dit zijn leden van het managementteam, afdelingshoofden, enkele staffunctionarissen en medewerkers die de informatie ten behoeve van de evaluatie hebben aangeleverd. Deze belanghebbenden, die directe toegang hebben tot de informatie uit de benchmark, zijn in staat om de eventueel aanwezige invloed terug te herleiden naar de benchmark. Andere belanghebbenden, zowel intern als extern, krijgen geen informatie of alleen een samenvatting. Vaker zijn zij ook niet op de hoogte van het feit dat deze informatie afkomstig is uit de benchmark en kunnen hierdoor de eventuele evaluatie-invloed niet of minder goed herleiden naar de benchmark. Alle belanghebbenden in dit onderzoek behoren tot een select gezelschap dat volledig toegang heeft tot de benchmark. Het zijn dus allen personen die betrokken zijn in het benchmarkproces en de resultaten ontvangen. Tevens zijn, ten behoeve van de vergelijkbaarheid, zo veel mogelijk gelijksoortige functionarissen geïnterviewd. Hierdoor is er minder diversiteit in perspectieven aanwezig. In de digitale versie van dit onderzoek zijn de interviewverslagen wegens geheimhouding niet in de bijlagen opgenomen. De interviewverslagen zijn wel ter beschikking gesteld aan de faculteit managementwetenschappen van de OU.

Het beperken van het onderzoek tot de motiverende en gedragsmatige veranderprocessen wordt ingegeven door het feit dat deze in dit onderzoek het meest voor de hand liggen. Beide veranderprocessen vereisen een 'trigger', een actie of gebeurtenis die voorafgaat aan het veranderproces. In dit onderzoek is de evaluatie deze trigger. Bij deze twee veranderprocessen kan derhalve duidelijk de link gelegd worden met de evaluatie, hetgeen een positief effect heeft op de betrouwbaarheid. Bij de generieke en cognitieve/affectieve veranderprocessen is het verband met de evaluatie minder duidelijk en kunnen andere factoren de oorzaak zijn van de veranderprocessen. De motiverende en gedragsmatige veranderprocessen focussen zich vooral op de korte en middellange termijn en kennen dus, in relatie tot de benchmark, een grotere mate van nauwkeurigheid dan de generieke en cognitieve/affectieve veranderprocessen. Deze laatste twee focussen zich met name op de middellange tot lange termijn.

4 RESULTATEN VAN HET ONDERZOEK

In dit hoofdstuk staan de onderzoeksresultaten centraal. Aan bod komen een verantwoording van de gehanteerde onderzoeksmethoden, de gegevensanalyse, de kwaliteit van de verzamelde gegevens en de kwaliteit van het proces van dataverzameling. Daarnaast wordt er een antwoord gegeven op de probleemstelling en de onderliggende deelvragen. Tot slot worden de beperkingen met betrekking tot de onderzoeksresultaten kort toegelicht.

Volgens Yin (2009) bestaan er vier generieke strategieën voor het analyseren van data, te weten:

1. gebaseerd op theoretische stellingen;
2. ontwikkelen van een casusomschrijving;
3. gebruik maken van zowel kwalitatieve- als kwantitatieve data;
4. het onderzoeken van tegengestelde verklaringen.

De eerste strategie, gebaseerd op theoretische stellingen, sluit het beste aan bij de ingebedde enkelvoudige gevalsstudie die in dit onderzoek wordt gebruikt. In deze gevalsstudie is namelijk een conceptueel model ontwikkeld op basis van een theoretisch kader. In dit model zijn stellingen benoemd en beschreven die teruggeleid kunnen worden naar de centrale onderzoeksvraag en de hieraan verbonden deelvragen en het uitgevoerde literatuuronderzoek. Vervolgens is dit conceptueel model en het onderliggende theoretisch kader gebruikt voor het opstellen van vragenlijsten. Door gebruik te maken van deze vragenlijsten bij de interviews is een bepaalde mate van volledigheid en uniformiteit geborgd waarbij gericht data is vergaard. Hierdoor was het mogelijk om andere, niet relevante, data buiten beschouwing te laten. Dit is feitelijk een voorbeeld van de uitvoering van de strategie van Yin (2009) met betrekking tot theoretische stellingen.

Dit onderzoek is een verklarend onderzoek. Het opbouwen van een verklaring (Yin, 2009) is logischerwijs de analysetechniek die hier het beste bij aansluit. Het uitgangspunt hierbij is het vormen van een theoretische verklaring. De bevindingen uit de casus vormen de basis voor deze theoretische verklaring.

4.1 Evaluatiesettingfactoren

Om de vijfde deelvraag van het onderzoek te kunnen beantwoorden, is achterhaald welke factoren, behorende tot de evaluatiesetting, volgens belanghebbende van belang zijn voor de benchmark. Deze informatie is achterhaald door eerst, op basis van literatuuronderzoek, vast te stellen welke factoren feitelijk behoren tot de evaluatiesetting. Vervolgens zijn deze factoren vertaald naar relevante interviewvragen om daarna met behulp van interviews te leiden tot de gewenste informatie. Triangulatie van databronnen was bijna niet mogelijk omdat het ervaringen en belevingen van de belanghebbenden betreft. Waar mogelijk zijn de antwoorden van belanghebbenden bekrachtigd door beschikbare documenten (triangulatie van databronnen) binnen de participerende organisaties en/of antwoorden van andere belanghebbenden binnen dezelfde organisatie. Er zijn in totaal dertien belanghebbenden geïnterviewd; zeven respondenten bij Brabant Water en zes respondenten bij WML. Het resultaat is

bepaald door het samenvoegen van de gegevens. Hierdoor worden de kwaliteit en betrouwbaarheid van de gegevens vergroot. In de digitale versie van dit onderzoek zijn de interviewverslagen wegens geheimhouding niet in de bijlagen opgenomen. De interviewverslagen zijn wel ter beschikking gesteld aan de faculteit managementwetenschappen van de OU.

Onderzocht is welke factoren in het kader van de evaluatiesetting van de benchmark een rol spelen bij WML en Brabant Water. Voor de factoren is gebruik gemaakt van het model dat is ontwikkeld door Cousins en Leithwood (Cousins & Leithwood, 1986). Zij identificeerden de volgende 6 factoren en classificeerden hen als behorende tot de evaluatiesetting: kwaliteit, geloofwaardigheid, relevantie, timing, resultaten en communicatie van de evaluatie.

Uit de interviews is gebleken dat bijna alle factoren een rol van betekenis spelen voor de benchmark. De geloofwaardigheid van de benchmark hangt sterk samen met de kwaliteit van de benchmark. Dit geldt ook voor de resultaten van de benchmark. De verschillende belanghebbenden geven aan dat zowel de geloofwaardigheid als de resultaten afhankelijk zijn van de kwaliteit van de benchmark. Het belang van alle drie deze factoren is groot bij zowel WML als Brabant Water, Brabant Water typeert de benchmark als zeer geloofwaardig en WML als geloofwaardig.

Op basis van analyse van de antwoorden met betrekking tot de factor 'relevantie' blijkt dat de benchmark voor WML als organisatie minder relevant is in vergelijking met Brabant Water. Bijna alle belanghebbenden binnen WML vinden de benchmark relevant voor henzelf maar niet voor WML. De achterliggende redenen voor dit antwoord hebben met name te maken met de tijd en moeite die WML, volgens belanghebbenden, steekt in de benchmark. Uitzondering hierop is de directeur van WML die de benchmark wel als relevant betitelt voor WML (zie bijlage 3, interview Ria Doedel). Bij Brabant Water is er bijna geen verschil tussen persoonlijke en organisatorische relevantie.

Bij Brabant Water wordt volgens de belanghebbenden voldoende gecommuniceerd over de benchmark en is deze communicatie in lijn met hun verwachtingen. De informatie wordt centraal beschikbaar gesteld en afgestemd op de verschillende doelgroepen. Volgens een van de belanghebbenden: "Iedereen moet weten wat wij met de benchmark willen bereiken, medewerkers moeten tot in de sleuf weten dat hij er is en dat erop gestuurd wordt" (zie bijlage 3, interview Leo de Waal). Bij WML vinden de meeste belanghebbenden, dat er onvoldoende wordt gecommuniceerd met betrekking tot de benchmark. Hierbij wordt er ook duidelijk minder waarde gehecht aan de communicatiekwaliteit in vergelijking met collega's bij Brabant Water. Enkele belanghebbenden geven aan dat dit met name wordt veroorzaakt door de in hun ogen beperkte relevantie.

De timing van de benchmark is volgens belanghebbenden van zowel WML als Brabant Water van ondergeschikt belang. Tijdens de interviews is een aantal keer gesteld dat het proces van continue verbeteren belangrijker is dan de resultaten zelf. Wel is het volgens de belanghebbenden van belang dat de benchmark een vaste plek in de planning en control-cyclus heeft. Bij Brabant Water wordt de benchmark ook gebruikt als check in de 'plan, do, check, act cyclus' (zie bijlage 3, interview Leo de Waal).

In het geval van WML is het antwoord op de vijfde deelvraag dat met name de met elkaar samenhangende factoren 'kwaliteit', 'geloofwaardigheid' en 'resultaten' van belang zijn voor de benchmark. Bij Brabant Water dienen hier de factoren 'relevantie' en 'communicatie' aan worden toegevoegd. Samenvattend ziet dit er zo uit:

Factor	WML	Brabant Water
Kwaliteit evaluatie	+	+
Geloofwaardigheid evaluatie	+	++
Resultaten evaluatie	+	+
Relevantie evaluatie	nvt	++
Communicatie evaluatie	nvt	+

Tabel 2. Resultaat evaluatiesettingfactoren

4.2 Besluitvorming en beleidsettingfactoren

Om de zesde deelvraag van het onderzoek te kunnen beantwoorden, is achterhaald welke factoren behorende tot de besluitvorming- en beleidsetting volgens belanghebbende van belang zijn voor de benchmark. Om deze informatie te achterhalen is dezelfde methodiek toegepast zoals toegelicht bij deelvraag 5 (zie eerste alinea 4.1).

Cousins en Leithwood hebben de volgende factoren geïdentificeerd en geclassificeerd onder de noemer besluitvorming- en beleidsetting: informatienoodzaak, besluitvormingskenmerken, politieke klimaat, concurrerende informatie, persoonlijke kenmerken en commitment met betrekking tot de evaluatie.

De belanghebbenden beschouwen de benchmark als waardevol en uniek en volgens hen bestaat er geen ander goed alternatief. De benchmark is uniek binnen de drinkwatersector, niet alleen in Nederland maar ook internationaal. De factor 'concurrerende informatie' speelt geen verdere rol van betekenis in dit onderzoek omdat er geen sprake is van invloed als gevolg van concurrerende informatie.

Alle belanghebbenden zijn positief over deelname aan de benchmark en willen de resultaten gebruiken in hun werk. De factor 'commitment' is van belang voor het succes van de benchmark maar speelt in deze casus geen rol van betekenis aangezien alle belanghebbenden hetzelfde hebben geantwoord.

De informatie die de benchmark oplevert is voor Brabant Water niet noodzakelijk. De belanghebbenden binnen Brabant Water vinden dat de informatienoodzaak een kleine rol van betekenis speelt voor de benchmark. De belanghebbenden binnen WML zijn hierin verdeeld. De helft vindt de informatie wel noodzakelijk en de andere helft niet. De belanghebbenden binnen WML zijn wel allen van mening dat de informatienoodzaak een rol van betekenis speelt voor de benchmark.

De besluitvormingskenmerken bij WML worden zeer verschillend getypeerd, namelijk: kaderstellend, incident gestuurd, integraal management, top down en moeizaam. Bij Brabant Water noemden belanghebbenden minder besluitvormingskenmerken, te weten: top down met een grote betrokkenheid

en consensus vanuit de organisatie. Zowel belanghebbenden binnen Brabant Water als WML vinden dat de besluitvormingskenmerken van belang zijn voor de benchmark.

Bij zowel Brabant Water als WML speelt de factor 'politiek' een rol in relatie tot de benchmark. Het is opvallend dat bij Brabant Water alle belanghebbenden eensgezind zijn over deze politieke factor maar bij WML niet. Daarnaast wordt bij WML zichtbaar dat de directie een groter belang hecht aan het politieke klimaat in vergelijking met andere belanghebbenden. Op basis van de analyse van de antwoorden valt tevens duidelijk af te leiden dat het politiek belang bij Brabant Water groter is dan bij WML. Bij Brabant Water wordt de benchmark bewust ingezet en gebruikt voor politieke doeleinden (zie bijlage 3, interview Guiljo van Nuland).

De invloed van persoonlijke kenmerken is volgens de belanghebbenden zeer gedifferentieerd. Enkele belanghebbenden geven aan dat deze geen rol van betekenis mogen spelen en stellen dat de benchmark onderdeel is van hun functie en verantwoordelijkheden. Andere belanghebbenden zien wel degelijk invloed van persoonlijke kenmerken. "Wanneer iemand weinig heeft met getallen dan is deze persoon ook minder genegen de benchmark te gebruiken" (zie bijlage 3 interview Erwin Peters). Tussen de antwoorden van belanghebbenden bij WML en Brabant Water doen zich hierin geen duidelijke verschillen voor.

In het geval van WML is het antwoord op de zesde deelvraag dat de factoren 'informatienoodzaak' en 'besluitvormingskenmerken' en in mindere mate de factoren 'politiek klimaat' en 'persoonlijke kenmerken' van belang zijn voor de benchmark. Bij Brabant Water zijn dit dezelfde factoren, maar het belang van het politieke klimaat is groter. Samenvattend ziet dit er zo uit:

Factor	WML	Brabant Water
Informatienoodzaak	++	++
Besluitvormingskenmerken	++	++
Politiek klimaat	+	++
Persoonlijke kenmerken	+	+

Tabel 3. Resultaat besluitvorming- en beleidsettingfactoren

4.3 Niveaus van invloed

Om de zevende deelvraag van dit onderzoek te kunnen beantwoorden is achterhaald welke verschillende niveaus van invloed door belanghebbenden bij WML en Brabant Water worden onderkend. Eerst is vastgesteld, op basis van literatuuronderzoek, welke niveaus van invloed er bestaan. Volgens Mark en Henry (Mark & Henry, 2003) zijn dit er drie, te weten individueel, interpersoonlijk en collectief. In dit onderzoek worden alleen de motiverende en gedragsmatige veranderprocessen belicht. De generieke en cognitieve affectieve veranderprocessen worden niet meegenomen. Vervolgens zijn deze niveaus van invloed vertaald naar relevante interviewvragen om daarna met behulp van interviews te leiden tot de gewenste informatie. Om deze informatie te achterhalen is dezelfde methodiek toegepast zoals

toegelicht bij deelvraag 5 (zie eerste alinea 4.1). Achtereenvolgens worden de resultaten per bedrijf toegelicht. Samenvattend worden de volgende niveaus van invloed onderzocht:

Individueel

- Verandering van houding (herzien en veranderen van mening);
- Verwerven van vaardigheden (evaluatie competentie);
- Verandering van gedrag (nieuwe concepten kunnen er toe leiden dat personen hun gedrag gaan veranderen).

Interpersoonlijk

- Veranderinstrument (resultaten gebruiken om anderen aan te zetten tot verandering);
- Sociale normen (resultaten leiden tot het bijstellen van een normenkader).

Collectief

- Beleid veranderen (beleid wordt aangepast als gevolg van de resultaten van de evaluatie);
- Diffusie (veranderingen in beleid en programma's als gevolg van de resultaten van de evaluatie bij organisaties die buiten het onderzoeksgebied van de evaluatie vallen).

4.3.1 Niveaus van invloed in het geval van WML

Bij WML zijn alle drie de niveaus van evaluatie-invloed aanwezig. De meeste belanghebbenden bij WML geven aan dat hun houding als gevolg van de benchmark is gewijzigd. Zo is een van de belanghebbenden (zie bijlage 3 interview Erwin Peters) zich meer extern gaan richten (focus op andere bedrijven). Twee belanghebbenden zijn kritischer naar het distributieproces gaan kijken, vooral als gevolg van de 'slechte scores' in de benchmark.

Geen van de geïnterviewde belanghebbenden heeft aangegeven nieuwe vaardigheden te hebben verworven. Er is wel een verandering in gedrag waarneembaar, maar deze is minimaal volgens de meeste belanghebbenden, met uitzondering van de directie. De volgende gedragsveranderingen zijn genoemd: nieuwsgieriger naar andere (inrichting van) processen, kritischer op het gebied van kosten, maken van strakkere afspraken en het beperken van de vrijblijvendheid van de benchmark. De genoemde punten zijn slechts voor enkele belanghebbenden van toepassing. Er is geen sprake van een collectieve gedragsverandering binnen de organisatie.

Alle belanghebbenden bij WML geven aan dat zij anderen, met behulp van de benchmark, hebben kunnen aanzetten tot verandering (interpersoonlijk). De invloed focust zich met name op het vergroten van het kostenbewustzijn bij medewerkers. De resultaten zijn goed vergelijkbaar en geven duidelijk de verschillen weer. De benchmark wordt actief gebruikt voor het aansturen van regio's en het aanzetten tot verandering. Een van de tastbare effecten is de daling in overwerk in een van de regio's (Bevindingen en aanbevelingen n.a.v. bedrijfsbezoek Brabant Water in relatie tot de benchmark 2009, 2009). Door meer in termen van kosten te denken is overwerk in deze regio afgenomen (zie bijlage 3 interview Erwin Peters). Geen van de belanghebbenden geeft aan als gevolg van de benchmark zijn of haar sociale normen te hebben gewijzigd.

Alle belanghebbenden geven aan dat het beleid van WML is veranderd als gevolg van de benchmark (collectief). In aanloop naar de recentelijk doorgevoerde reorganisatie is de benchmark herhaaldelijk genoemd maar concrete invloed kan men niet direct toewijzen. Daarnaast zou de benchmark ook aan de basis hebben gestaan van de Operational Excellence-strategie. Een van de hierin opgenomen doelstellingen is een vierde plaats (best of the rest) in de benchmark op operationele kosten (WML Presentatie RvC MN 2012-2016, 2011). Deze invloed is echter niet tastbaar of moeilijk aan te tonen. Een concreet voorbeeld echter is het recentelijk genomen besluit om de controlefrequentie van afsluiters aan te passen. Dit besluit is genomen op basis van een bezoek aan Brabant Water en is direct te herleiden naar de benchmark (zie bijlage 3 interview Ria Doedel).

De benchmark heeft ook invloed gehad buiten de Nederlandse drinkwatersector. Wegens het aangetoonde succes binnen de Nederlandse drinkwatersector is er nu ook een benchmark geïntroduceerd bij de Waterschappen met als doel een kostenreductie te realiseren.

4.3.2 Niveaus van invloed in het geval van Brabant Water

Bij Brabant Water zijn ook de drie verschillende niveaus van evaluatie-invloed aanwezig. De helft van de belanghebbenden geeft aan zijn gedrag of houding niet te hebben veranderd. De andere helft is bewuster en scherper geworden met betrekking tot kosten en processen. De coördinator van de benchmark bij Brabant Water is verantwoordelijk voor het presenteren van de benchmarkresultaten. Als gevolg van de benchmark heeft hij presentatievaardigheden verworven (zie bijlage 3 interview Henk de Boer). Geen andere belanghebbende geeft aan, als gevolg van de benchmark, nieuwe vaardigheden te hebben verworven. Twee van de belanghebbenden bij Brabant Water geven echter wel aan hun gedrag te hebben veranderd. Een belanghebbende is meer kostenbewust geworden en meer gaan analyseren, de ander is zijn afdeling anders gaan aansturen door aspecten van de benchmark mee te nemen in zijn besprekingen.

Het collectief beleid van Brabant Water heeft veel invloed ondervonden van de benchmark. Alle belanghebbenden noemen hierbij de doelstelling dat Brabant Water in de top 3 wil staan van niet één maar alle benchmarkprocessen. De distributieafdelingen hebben zelfs een eigen visie/beleidsdocument opgesteld 'krachtig op koers 2016' waarin zij aangeven in 2016 voor wat betreft het distributieproces op één te willen staan in de benchmark (zie bijlage 3 interview John de Schepper en document krachtig op koers 2016, 2011). Recente keuzes met betrekking tot het verbeteren van de waterkwaliteit en het voornemen een investeringsprogramma van meer dan 100 miljoen te effectueren is ook ingegeven door de slechte positie in de benchmark op waterkwaliteit. Op basis hiervan is een afweging gemaakt: "hoeveel extra kosten kunnen wij als Brabant Water in het productieproces maken rekeninghoudend met top 3 doelstelling om ons water kwalitatief te verbeteren?" (zie bijlage 3 interview Leo de Waal). Kosten en kwaliteit zijn speerpunten in het beleid van Brabant Water die volgens belanghebbenden duidelijk herleidbaar zijn naar de benchmark. Afwegingen worden integraler gemaakt, er wordt rekening gehouden met gevolgen op andere processen. Hierdoor is het beleid van Brabant Water volgens belanghebbenden evenwichtiger geworden, meer in balans. Een belanghebbende binnen Brabant Water verwoordt het als volgt: "De benchmark leeft, wordt beleefd en binnen Brabant Water bestaat er een sense of urgency waar het de benchmark betreft" (zie bijlage 3 interview Guiljo van Nuland). Een andere

belanghebbende verwoordt het als volgt: “Wij zetten anderen aan tot verandering door kennis betreffende de werking van de benchmark over te brengen” (zie bijlage 3 interview Jos Bierens). Ook het creëren van de nieuwe functie ‘project-stuurder’ wordt beschouwd als een direct gevolg van de benchmark. Een ander voorbeeld is het bewust in eigen dienst behouden van de monsternemers (Analyse benchmarkproces Bewaken waterkwaliteit PO-BWK, 2007).

4.3.3 Verschillen en overeenkomsten WML en Brabant Water

In het geval van WML en Brabant Water is het antwoord op de zevende deelvraag dat alle drie de niveaus van evaluatie-invloed worden onderkend. Bij WML ligt de nadruk vooral op individuele en interpersoonlijke invloed, zoals het kritisch zijn op kosten en een meer externe focus. Belanghebbenden bij WML onderkennen ook collectieve invloed maar deze is nauwelijks concreet en erg beperkt. Bij Brabant Water ligt de nadruk zowel op individuele invloed als op collectieve invloed. Alle belanghebbenden noemden de top 3-doelstelling binnen de benchmark en dat deze doelstelling sterk wordt uitgedragen binnen alle lagen van de organisatie. De benchmark is een integraal onderdeel in de bedrijfsvoering van Brabant Water. Alle belangrijke beleidsmatige keuzes worden doorvertaald naar benchmarkeffecten en als zodanig ook meegenomen in de besluitvorming. Zie tabel 4:

Factor	WML	Brabant Water
Individueel	++	++
Interpersoonlijk	++	+
Collectief	+	++

Tabel 4. Resultaat evaluatie-invloed

4.4 Typen verandermechanismen

Het antwoord op de achtste deelvraag (welke typen verandermechanismen liggen aan de basis van de totstandkoming van de evaluatie-invloed in het geval van WML en Brabant Water?) is terug te vinden in de literatuur. Mark en Henry (Mark & Henry, 2004) hebben vier verschillende typen verandermechanismen geïdentificeerd. Door deze in een breder perspectief te plaatsen hebben ze een model ontwikkeld (zie fig. 3, p 19) waarin duidelijk de link te zien is tussen de drie gestelde niveaus van evaluatieresultaten en de onderliggende verandermechanismen.

Bij individuele invloed wordt, een verandering van houding, gedrag of het verwerven van nieuwe- of anders inzetten van bestaande vaardigheden, ingegeven door een verandermechanisme op het gebied van gedrag. Bij interpersoonlijke invloed wordt een verandering in normen toegewezen aan een motivationeel verandermechanisme. Het beïnvloeden van het gedrag van een ander wordt toegewezen aan een verandermechanisme op het gebied van gedrag. Bij collectieve invloed wordt een beleid- of programmaverandering toegewezen aan een gedragsmatig verandermechanisme. Dit is ook het geval voor de niet bedoelde collectieve evaluatie-invloed buiten de Nederlandse drinkwatersector.

Het antwoord op de achtste deelvraag is dan ook dat bij zowel WML als Brabant Water de gedragsmatige verandermechanismen aan de basis hebben gestaan bij de totstandkoming van de evaluatie-invloed.

4.5 Invloed factoren evaluatiesetting

Het antwoord op de negende deelvraag (op welke manieren verklaren de factoren in het kader van de evaluatiesetting van de benchmark de invloed van de benchmark in het geval van Brabant Water en WML?) wordt achterhaald door de antwoorden van de vijfde en zevende deelvraag te analyseren.

Volgens de belanghebbenden bij zowel WML als Brabant Water heeft de benchmark invloed. Op basis van de interviews kan geconcludeerd worden dat de invloed van de benchmark bij Brabant Water groter is dan bij WML. In de digitale versie van dit onderzoek zijn de interviewverslagen wegens geheimhouding niet in de bijlagen opgenomen. De interviewverslagen zijn wel ter beschikking gesteld aan de faculteit managementwetenschappen van de OU. Achtereenvolgens worden de resultaten per bedrijf toegelicht.

4.5.1 Invloed factoren evaluatiesetting in het geval van WML

Bij WML wordt de kwaliteit van de benchmark op hoofdlijnen licht voldoende bevonden. Op detailniveau worden bij de gegevens, die door WML zelf worden aangeleverd, door belanghebbenden vraagtekens geplaatst en is de kwaliteit volgens hen niet voldoende. Dit heeft zijn weerslag op de resultaten en de geloofwaardigheid van de benchmark. Op hoofdlijnen vinden belanghebbenden bij WML de benchmark geloofwaardig en geeft het de juiste resultaten weer, maar op detailniveau zijn de gegevens niet betrouwbaar en daardoor, volgens de belanghebbenden, ook niet geloofwaardig. Dit onderscheid tussen hoofdlijnen en details verklaart ook de factor relevantie. Op directieniveau is de relevantie hoger dan op afdelingsniveau. De benchmark is betrouwbaar op bedrijfsniveau en dus relevant voor de directie. Op afdelingsniveau zijn belanghebbenden minder tevreden met de resultaten en de betrouwbaarheid van de benchmark. Logischerwijs is de relevantie ook lager. Doordat volgens belanghebbenden alleen het hoognodige wordt gecommuniceerd, beperkt de communicatie zich vooral tot de hoofdlijnen en is met name gericht op financiële prestaties. Daarnaast worden de resultaten maar aan een beperkt aantal personen gecommuniceerd. De timing van de benchmark speelt geen rol van betekenis en is geen verklarende factor voor de invloed van de benchmark.

De invloed van de benchmark bij WML wordt vooral verklaard door de factoren 'relevantie' en 'communicatie'. De benchmark wordt door verschillende belanghebbenden bij WML niet als relevant beschouwd. Dit wordt versterkt door het beeld dat belanghebbenden hebben dat de benchmark op hoofdlijnen wel betrouwbaar is maar op detailniveau niet. Belanghebbenden lijken hier eigenlijk voorbij te gaan aan het feit dat deze details door WML zelf worden aangeleverd. Belanghebbenden worden onvoldoende geïnformeerd betreffende de achtergrond en onderbouwing van de opgevoerde data in de benchmark. Dit kan direct herleid worden naar de factor communicatie. Veel belanghebbenden hebben aangegeven niet te weten wanneer het proces van dataverzameling voor de benchmark start. De factoren geloofwaardigheid, kwaliteit en resultaten zijn sterk afhankelijk van de kwaliteit van invoer van gegevens. Bij een beperkte relevantie zal men sneller genoegen nemen met minder kwaliteit. Wanneer de benchmark een hogere relevantie heeft, zal de kwaliteit van de input en de hiermee samenhangende

nauwkeurigheid en herkenbaarheid vanzelfsprekend groter zijn. De factor communicatie speelt hierin ook een rol. Binnen WML wordt maar beperkt aan een selecte groep medewerkers over de benchmark gecommuniceerd. Hierdoor nemen de relevantie en invloed van de benchmark niet toe. De beperkte communicatie zorgt er voor dat de huidige situatie, dat wil zeggen de gelimiteerde evaluatie-invloed, blijft voortbestaan.

4.5.2 Invloed factoren evaluatiesetting in het geval van Brabant Water

Bij Brabant Water voldoet de benchmark volgens belanghebbenden op hoofdlijnen aan de kwaliteits-eisen. Ook hier worden door belanghebbenden vraagtekens geplaatst op detailniveau en is de kwaliteit niet voldoende. Deze vraagtekens hebben echter niet betrekking op de eigen informatie maar op de opgevoerde datasets van andere bedrijven. In lijn met WML heeft ook dit gevolgen voor de resultaten en de geloofwaardigheid van de benchmark. Op hoofdlijnen (hoofdproces) vindt men de benchmark zeer geloofwaardig en geeft de benchmark ook de juiste resultaten weer. Op detailniveau (deelproces) zijn de resultaten van Brabant Water zeer betrouwbaar en geloofwaardig maar de detailinformatie van de andere deelnemende bedrijven wordt met minder vertrouwen gezien.

Het officiële externe eindrapport van de benchmark wordt op hoofdlijnen gepresenteerd en bevat geen details. Deze rapportage geeft de juiste resultaten weer en is geloofwaardig. Dit onderscheid tussen extern en intern gebruik verklaart ook de factor relevantie. Alle belanghebbenden bij Brabant Water beschouwen de benchmark als zeer relevant. Hier is geen onderscheid tussen de verschillende rollen en functies van de geïnterviewde belanghebbenden. Tevens wordt er binnen Brabant Water veelvuldig en op verschillende niveaus over de benchmark gecommuniceerd. Alle afdelingen en medewerkers worden geïnformeerd over de benchmark. Iedereen dient op de hoogte te zijn van het bestaan van de benchmark en hetgeen wat er mee beoogd wordt. Ook extern wordt de benchmark veelvuldig ingezet bij communicatie met belangrijke stakeholders.

De invloed van de benchmark bij Brabant Water wordt dan ook vooral verklaard door de factoren 'relevantie', 'geloofwaardigheid' en 'communicatie'. Alle belanghebbenden bij Brabant Water betitelen de benchmark als zeer geloofwaardig. Hierbij moet echter wel rekening gehouden worden met het hierboven beschreven onderscheid tussen hoofd- en detailniveau en tussen intern en extern. De benchmark wordt daarnaast als zeer relevant gezien. De doelstelling om top 3 te willen scoren in alle benchmarkprocessen is één op één gekoppeld aan de benchmarkresultaten. Dit is een strategische doelstelling die door alle belanghebbenden gedragen en gecommuniceerd wordt. De benchmark wordt intern vooral ingezet als verbeterinstrument. Medewerkers worden uitgedaagd om beter en efficiënter te werken waarbij de benchmark als leidraad wordt gebruikt. De communicatie betreffende de benchmark beperkt zich dan ook niet tot één maal per jaar een presentatie. Elk kwartaal worden de behaalde resultaten doorvertaald naar de benchmark en wordt een tussenstand gerapporteerd op afdelingsniveau. Afdelingen concurreren onderling met elkaar maar leren ook van elkaar. In de communicatie wordt de benchmark vooral getypeerd als verbeterinstrument in plaats van vergelijking- of afrekeninstrument. Hierdoor verklaart ook de factor 'communicatie' een belangrijk aspect van de evaluatie-invloed van de benchmark bij Brabant Water.

Het antwoord op de negende deelvraag kan als volgt worden samengevat:

Invloed Factor	WML	Brabant Water
Relevantie	+	++
Communicatie	+	++
Geloofwaardigheid	nvt	++

Tabel 5. Verklaren invloed evaluatiesetting factoren

4.6 Invloed factoren besluitvorming- en beleidsetting

Het antwoord op de tiende deelvraag (op welke manieren verklaren de factoren in het kader van de besluitvorming- en beleidsetting van de benchmark de invloed van de benchmark in het geval van Brabant Water en WML?) wordt achterhaald door de antwoorden van de zesde en zevende deelvraag te analyseren. Achtereenvolgens worden de resultaten per bedrijf toegelicht.

4.6.1 Invloed factoren besluitvorming- en beleidsetting in het geval van WML

De belanghebbenden bij WML zijn van mening dat de informatie die de benchmark bevat niet noodzakelijk is voor de bedrijfsvoering van WML. De informatie is echter wel wenselijk en van toegevoegde waarde voor het benoemen van verbeterpotentieel. Belanghebbenden menen dat zonder de benchmark uiteindelijk dezelfde resultaten behaald worden, maar dit traject zou dan meer tijd kosten. De benchmark wordt dus vooral gezien als een katalysator die het veranderproces versneld heeft. Over de besluitvormingskenmerken zijn de meningen binnen WML verdeeld. Vier belanghebbenden vinden deze factor erg belangrijk voor de benchmark en twee belanghebbenden juist niet. Dit is ook terug te zien in de typering van de besluitvormingskenmerken. De antwoorden 'top down' en 'mag meer top down zijn', 'integraal management', 'incident gestuurd', 'chaotisch' en 'elk besluit is de start van een nieuwe discussie' zijn allemaal genoemd (zie bijlage 3). Wel is men het er over eens dat een goede eenduidige besluitvorming positief zou doorwerken op de mogelijke invloed van de benchmark.

Eenzelfde verdeeldheid is ook te zien bij de factor 'politiek klimaat'. Twee belanghebbenden geven aan dat de politiek geen rol speelt bij de benchmark, de overige vier geven aan dat dit wel het geval is. Waarbij één belanghebbende aangeeft dat deze rol nog belangrijker gaat worden en een ander juist aangeeft dat de rol in de toekomst kleiner gaat worden. Volgens de belanghebbenden sluit de benchmark wel goed aan bij de wensen van de interne politiek, maar wordt dit niet actief uitgedragen of hardop uitgesproken. Het externe politieke klimaat is voor de meeste belanghebbenden onbekend. In hun functie geven zij aan weinig van doen te hebben met stakeholders buiten WML. Recente ontwikkelingen op politiek vlak geven wel aan dat de rol van de overheid binnen de Nederlandse drinkwatersector groter gaat worden.

Aangezien de benchmark uniek is binnen de Nederlandse drinkwatersector speelt de factor 'concurrerende informatie' geen rol van betekenis.

Volgens de meeste belanghebbenden bij WML spelen, in relatie tot de benchmark, de persoonlijke kenmerken geen rol van betekenis. Natuurlijk is het van belang dat een persoon analytisch is onderlegd, maar men koppelt dit kenmerk meer aan de functie dan aan de persoon. Bepaalde competenties zijn noodzakelijk om op een bepaald niveau te kunnen opereren. Alle belanghebbenden bevinden zich op

een niveau binnen de organisatie waarbinnen bepaalde competenties vereist zijn. Alle belanghebbenden bij WML zijn hoog opgeleid.

De factor 'commitment' is van belang voor de benchmark. Zonder commitment is er geen draagvlak en zouden er ook geen resultaten en/of invloed zijn. Alle belanghebbenden onderschrijven het belang van commitment en ervaren dit ook. Wel geeft één belanghebbende aan steeds minder positief te worden. (zie bijlage 3 interview Ruud Vliegen) De benchmark begint steeds meer te verworden tot een verplichting waardoor het commitment afneemt.

Kortom, de invloed van de benchmark bij WML wordt vooral verklaard door de factoren 'besluitvormingskenmerken', 'politiek klimaat', 'persoonlijke kenmerken' en 'commitment'. De besluitvorming binnen WML wordt zo divers beschreven dat niet gesproken kan worden van een eenduidig of gestructureerd proces. Door een lager draagvlak voor besluiten is de mogelijke invloed van de benchmark als gevolg van directiebesluiten laag. De benchmark sluit echter wel goed aan bij het interne politieke klimaat (kostenbeheersing). Belanghebbenden met een positieve attitude richting de benchmark zullen deze dan ook gebruiken om hier bij aan te sluiten. Dit zijn personen die meer affiniteit hebben met cijfers en de benchmarkresultaten begrijpen. De invloed van de benchmark bij WML komt vooral tot stand door individuen die de benchmark begrijpen en waarderen. Ze zien het potentieel van de benchmark en gebruiken deze voor het verbeteren van hun proces. De invloed vindt vooral plaats op afdelingsniveau en niet op bedrijfsniveau en is incidenteel van aard in plaats van structureel.

4.6.2 Invloed factoren besluitvorming- en beleidsetting in het geval van Brabant Water

Ook de belanghebbenden bij Brabant Water zijn van mening dat de informatie die de benchmark bevat niet noodzakelijk is voor de bedrijfsvoering. Letterlijk is gezegd dat "de drive voor presteren en verbeteren niet wordt ingegeven door de benchmark" (zie bijlage 3 interview Kirk den Otter). Er is echter wel duidelijk behoefte aan de benchmark, maar dan vooral als bevestiging dat geleverde inspanningen hebben geleid tot het gewenste resultaat. De factor 'informatienoodzaak' verklaart niet de aanwezige invloed. In tegenstelling tot WML zijn alle belanghebbenden bij Brabant Water het unaniem eens over de besluitvormingskenmerken. Deze worden door iedereen getypeerd als top down, maar met veel inspraak en betrokkenheid vanuit de organisatie. Genomen besluiten kunnen rekenen op veel draagvlak en worden gemakkelijk geaccepteerd. De belanghebbenden zijn wel verdeeld over het belang van de besluitvormingskenmerken voor de benchmark. De belanghebbenden behorende tot de directie kennen een minder belang toe aan de besluitvormingskenmerken dan de overige belanghebbenden, die de besluitvormingskenmerken zeer waarderen en niet vanzelfsprekend vinden. Besluiten vanuit de directie kunnen dus op veel draagvlak rekenen binnen de organisatie. De factor besluitvormingskenmerken is daarom bij Brabant Water een belangrijke verklaring voor de aanwezige collectieve invloed. De top 3 halen in alle benchmarkprocessen is een van de belangrijkste doelstellingen van Brabant Water. Deze doelstelling is op directieniveau geformuleerd en wordt gedragen door de organisatie.

Ook met betrekking tot de factor 'politiek klimaat' zijn belanghebbenden bij Brabant Water eensgezind. Alle belanghebbenden geven aan dat de politiek een rol van betekenis speelt. De benchmark is een officieel rapport dat extern gepubliceerd wordt. Ook is veelvuldig aangegeven dat het ontstaan van de

benchmark een politieke lading heeft gehad. De benchmark wordt veelvuldig extern gebruikt en is daardoor ook intern van belang. De directie wil aan de belangrijkste stakeholders namelijk laten zien dat Brabant Water in vergelijking met andere drinkwaterbedrijven goed presteert en welke verbeterambities zij wil realiseren. Ook de top 3 doelstelling is extern benoemd en bekrachtigd. De keuze om de benchmark veelvuldig extern in te zetten, verklaart een groot deel van de aanwezige invloed.

Aangezien de benchmark uniek is binnen de Nederlandse drinkwatersector speelt de factor 'concurrerende informatie' geen rol van betekenis.

Bij Brabant Water verschillen de belanghebbenden van mening over het belang van de persoonlijke kenmerken voor de benchmark. Enkelen geven aan dat er een belang is en anderen juist niet. Zo stelt de een dat het van belang is dat een persoon analytisch is onderlegd en 'iets met resultaten heeft' (zie bijlage 3 interview John de Schepper), anderen belanghebbenden geven aan dat dit niet zozeer een persoonlijk kenmerk is maar dat dit hoort bij de functie. Een ondersteunende functie vereist andere competenties dan een operationele functie. De juiste persoon op de juiste plek met de juiste ondersteuning is belangrijk. Het resultaat is pas optimaal wanneer aan deze voorwaarden voldaan wordt. Alle belanghebbenden bij Brabant Water zijn hoog opgeleid. Commitment is van belang voor de benchmark. In lijn met WML geven ook de belanghebbenden bij Brabant Water aan dat zonder commitment er geen draagvlak voor de benchmark kan zijn en dus ook geen resultaten en invloed. Alle belanghebbenden geven aan positief te zijn betreffende de benchmark en het gebruik ervan.

Kortom, de invloed van de benchmark bij Brabant Water wordt vooral verklaard door de factoren 'besluitvormingskenmerken', 'politiek klimaat', 'persoonlijke kenmerken' en 'commitment'. De directie gelooft in de benchmark en heeft de benchmark tot een belangrijke doelstelling (top 3) gemaakt. Dit is in grote lijnen het gevolg van het commitment van enkele personen binnen Brabant Water naar de benchmark toe. Niet alleen committeert Brabant Water zich intern aan de benchmark, ook extern worden doelstellingen gecommuniceerd en vastgelegd. Hoewel het initiatief en de besluitvorming met name bij de directie ligt, wordt de benchmark toch enorm beleefd binnen de organisatie. De besluitvormingskenmerken hebben een belangrijke rol in dit proces. De collectieve invloed bij Brabant Water wordt met name veroorzaakt door de top 3 doelstelling. Veel besluiten die worden genomen zijn gebaseerd op deze doelstelling.

Het antwoord op de tiende deelvraag kan als volgt worden samengevat:

Invloed Factor	WML	Brabant Water
Besluitvormingskenmerken	+	++
Politiek klimaat	++	++
Persoonlijke kenmerken	++	+
Commitment	+	++

Tabel 6. Verklaren invloed besluitvorming- en beleidsettingfactoren

4.7 Verklaring verschil van invloed

Het antwoord op de elfde deelvraag (op welke manier kan het verschil van invloed van de evaluatie verklaard worden door de evaluatiesetting of besluitvorming- en beleidsetting factoren?) wordt achterhaald door de antwoorden van de negende en tiende deelvraag te analyseren.

Zowel de factoren behorende tot de evaluatiesetting als de factoren behorende tot de besluitvorming- en beleidsetting verklaren het verschil van invloed in het geval van WML en Brabant Water. Hoe geloofwaardig, relevant, tijdig en kwalitatief een evaluatie ook is, zonder de factoren behorende bij de besluitvorming- en beleidsetting zal een evaluatie nooit veel invloed kunnen uitoefenen.

De evaluatie-invloed bij WML is incidenteel van aard en met name ingegeven door individuen. De feitelijke invloed is beperkt en heeft vooral betrekking op het creëren van meer kostenbewustzijn. Op collectief niveau is het moeilijk om invloed specifiek te benoemen. Vaak worden keuzes wel in verband gebracht met de benchmark, maar kan dit nauwelijks concreet worden benoemd of bewezen. De evaluatie-invloed bij Brabant Water is structureel van aard. Het beperkt zich niet tot kostenbewustzijn maar betreft ook kennis over de processen. Op collectief niveau is de invloed groot. De benchmark heeft, mede als gevolg van de top 3 doelstelling, veelvuldig aan de basis gestaan van grote complexe besluiten. De benchmark wordt door Brabant Water structureel voor diverse doeleinden ingezet. De benchmark is veel relevanter voor Brabant Water dan voor WML. Als gevolg daarvan wordt er ook meer en op een andere manier over de benchmark binnen Brabant Water gecommuniceerd. Het verschil in relevantie en de hieraan gekoppeld communicatie, verklaart voor een groot deel het verschil in invloed. Bij een eenduidige, geaccepteerde besluitvorming is het gemakkelijker om structurele veranderingen als gevolg van de benchmark te realiseren. Bij Brabant Water verloopt dit proces makkelijker dan bij WML. Hierdoor is de evaluatie-invloed bij Brabant Water ook groter dan bij WML. Dit verklaart ook waarom de invloed bij WML incidenteel van aard is en bij Brabant Water meer gestructureerd en integraal.

Het verschil van invloed van de benchmark bij WML en Brabant Water wordt dus in grote mate verklaard door de factoren 'relevantie' en 'commitment' en in mindere mate door de factor 'besluitvormingskenmerken'. De relevantie van de evaluatie kan weer herleid worden naar de factor 'commitment'. De directeur van Brabant Water geeft aan dat hij de benchmark ontzettend belangrijk vindt en dat hij er voor zorgt dat zijn directieleden dit ook zo vinden. Hij is duidelijk zeer gecommitteerd aan de benchmark en verwacht dit ook bij anderen binnen Brabant Water. Op basis hiervan kan worden geconcludeerd dat in dit onderzoek de factoren behorende bij de besluitvorming- en beleidsetting van groter belang zijn voor de evaluatie-invloed dan de factoren behorende bij de evaluatiesetting.

5 CONCLUSIES, DISCUSSIE EN AANBEVELINGEN

In dit hoofdstuk komen de conclusies, discussie en aanbevelingen aan de orde. Op een synthetiserende wijze worden de belangrijkste resultaten van het onderzoek toegelicht. In de discussie wordt gereflecteerd op de betrouwbaarheid en validiteit van de gebruikte onderzoeksmethoden. Tevens wordt duidelijk gemaakt hoe de onderzoeksresultaten bijdragen aan de bestaande wetenschappelijke kennis/literatuur. Voorts worden de resultaten geanalyseerd ten opzichte van de theoretische bevindingen zoals deze aan de orde komen in hoofdstuk 2. Het hoofdstuk wordt afgesloten met aanbevelingen, waaronder relevante praktische implicaties, die uit het onderzoek volgen en de wijze waarop de resultaten en beperkingen van dit onderzoek aanleiding kunnen geven tot vervolgonderzoek.

5.1 Conclusies van het onderzoek

Voor het formuleren van de conclusie van het onderzoek zijn de centrale vraag, de deelvragen en de antwoorden op deze vragen van belang. De centrale vraag van dit onderzoek luidt als volgt: In welke mate verklaren de evaluatiesettingfactoren en de besluitvorming- beleidsettingfactoren de aanwezigheid van evaluatie-invloed in het geval van WML en Brabant Water? Om deze vraag te kunnen beantwoorden zijn elf deelvragen geformuleerd. De antwoorden op deze deelvragen worden in de onderliggende alinea's toegelicht. Deze paragraaf wordt afgesloten met het antwoord op de centrale vraag en de conclusie van het onderzoek.

Deelvraag 1. Wat wordt in wetenschappelijk onderzoek verstaan onder evaluatie-invloed?

Evaluatie-invloed verwijst naar het totale effect van een evaluatie. Het onderzoeksgebied 'evaluatie' heeft zich de afgelopen decennia verbreed door niet alleen meer te focussen op benutting maar het accent te verleggen naar de invloed van evaluaties. De term evaluatiebenutting is te gelimiteerd en dekt onvoldoende de lading wanneer men wil onderzoeken wat de veranderingen zijn die plaatsvinden als gevolg van een evaluatie.

Deelvraag 2. Welke verschillende niveaus van evaluatie-invloed kunnen worden onderscheiden?

Mark en Henry (Mark & Henry, 2003) hebben een model ontwikkeld waarin zij drie verschillende niveaus van evaluatie-invloed identificeren. Te weten individueel, interpersoonlijk en collectief. Op het individuele niveau benoemen Mark en Henry zes mogelijke uitkomsten, ook wel verandermechanismen genoemd, namelijk: verandering van houding, saillantie, uitwerking, priming, verwerven van vaardigheden en verandering van gedrag. Op het interpersoonlijke niveau identificeren Mark en Henry vijf mechanismen die veranderingen in houding en gedrag bij anderen te weeg kunnen brengen, namelijk: rechtvaardiging, overreding, veranderinstrument, sociale normen en de invloed van de opinie van een minderheid. Het collectieve niveau refereert naar de invloed van evaluatie op verschillende individuen en organisaties tegelijkertijd. Mark en Henry identificeren vier mechanismen, namelijk: agendabepaling, beleid georiënteerd leervermogen, beleid veranderen en diffusie.

Deelvraag 3. Welke typen verandermechanismen kunnen worden onderscheiden die de totstandkoming van evaluatie-invloed beïnvloeden?

Mark en Henry (Mark & Henry, 2004) hebben op basis van hun model verschillende verandermechanismen beschreven, direct verbonden aan de drie niveaus van invloed. Vervolgens hebben zij de geïdentificeerde mechanismen geclassificeerd en gecategoriseerd in vier verschillende typen. Te weten: generieke, cognitieve en affectieve, motiverende en gedragsmatige verandermechanismen. Door aandacht te schenken aan deze mechanismen en hun impact op de processen kunnen in de toekomst evaluaties beter gepland worden. Beoordelaars kunnen beter anticiperen op de processen die als gevolg van de evaluatie in gang worden gezet. Dit kunnen meerdere processen tegelijkertijd zijn waarbij processen ook getriggerd kunnen worden door andere processen (niet de evaluatie zelf). Elk individueel proces kan een korte termijn, intermediaire of lange termijn evaluatieresultaat zijn in de vaak complexe weg naar social betterment.

Deelvraag 4. Welke factoren kunnen de benutting van evaluaties beïnvloeden?

Bober en Bartlett (Bober & Bartlett, 2004) hebben als basis voor hun onderzoek het conceptuele model gebruikt dat is ontwikkeld door Cousins en Leithwood (Cousins & Leithwood, 1986). Ze kozen dit model vanwege de volledigheid en toepasselijkheid voor het onderzoeken van factoren en andere aspecten met betrekking tot de benutting van evaluaties. In dit model zijn de factoren benoemd en toegewezen aan twee categorieën. In de categorie 'evaluatiesetting' kunnen de volgende invloedrijke factoren benoemd worden: de kwaliteit van de evaluatie, de geloofwaardigheid van de evaluatie, de relevantie van de evaluatie, de communicatiekwaliteit van de evaluatie, de resultaten van de evaluatie en de timing van de evaluatie. In de categorie 'besluitvorming- en beleidsetting' kunnen de volgende factoren benoemd worden: de informatienoodzaak, de besluitvormingskenmerken van de organisatie, het politieke klimaat, de concurrerende informatie, de persoonlijke kenmerken van gebruiker en het commitment naar de evaluatie toe.

De antwoorden op de eerste 4 deelvragen vormen het theoretisch kader van het verklarende onderzoek en dienen als uitgangspunt voor het beantwoorden van de andere deelvragen.

Deelvraag 5. Welke factoren in het kader van de evaluatiesetting van de benchmark spelen een rol in het geval van WML en Brabant Water?

In het geval van WML spelen met name de met elkaar samenhangende factoren 'kwaliteit', 'geloofwaardigheid' en 'resultaten' een rol van betekenis. Bij Brabant Water zijn dit eveneens de met elkaar samenhangende factoren 'kwaliteit', 'geloofwaardigheid' en 'resultaten' maar hieraan kunnen nog de factoren 'relevantie' en 'communicatie' worden toegevoegd.

Deelvraag 6. Welke factoren in het kader van de besluitvorming- en beleidsetting van de benchmark spelen een rol in het geval van WML en Brabant Water?

In het geval van WML zijn voornamelijk de factoren 'informatienoodzaak' en 'besluitvormingskenmerken' en in mindere mate de factoren 'politicus klimaat' en 'persoonlijke kenmerken' van belang voor de Benchmark. Bij Brabant Water zijn dit dezelfde factoren, maar het belang van het politieke klimaat is groter.

Deelvraag 7. Welke verschillende niveaus van evaluatie-invloed worden, als gevolg van de benchmark, onderkend door de belanghebbenden in het geval van WML en Brabant Water?

In het geval van WML en Brabant Water worden alle drie de niveaus van evaluatie-invloed onderkend. Bij WML ligt de nadruk vooral op individuele en interpersoonlijke invloed, zoals het kritisch zijn op kosten en een meer externe focus. Belanghebbenden bij WML onderkennen ook collectieve invloed maar deze is nauwelijks concreet en zeer beperkt. Bij Brabant Water ligt de nadruk op individuele maar vooral op collectieve invloed. Alle belanghebbenden noemden de top 3-doelstelling binnen de benchmark en dat deze doelstelling sterk wordt uitgedragen binnen alle lagen van de organisatie. De benchmark is een integraal onderdeel in de bedrijfsvoering van Brabant Water. Alle belangrijke beleidsmatige keuzes worden doorvertaald naar benchmarkeffecten en als zodanig ook meegenomen in de besluitvorming.

Deelvraag 8. Welke typen verandermechanismen liggen aan de basis van de totstandkoming van deze evaluatie-invloed in het geval van WML en Brabant Water?

De gedragsmatige verandermechanismen hebben bij beide organisaties aan de basis gestaan bij de totstandkoming van de aanwezige evaluatie-invloed.

Deelvraag 9. Op welke manier verklaren de factoren in het kader van de evaluatiesetting de invloed van de Benchmark in het geval van WML en Brabant Water?

De invloed van de benchmark bij WML wordt vooral verklaard door de factoren 'relevantie' en 'communicatie'. De benchmark wordt door verschillende belanghebbenden bij WML niet als relevant beschouwd. Dit wordt versterkt door het beeld dat belanghebbenden hebben dat de benchmark op hoofdlijnen wel betrouwbaar is maar op detailniveau niet. Belanghebbenden worden onvoldoende geïnformeerd betreffende de achtergrond en onderbouwing van de opgevoerde data in de benchmark. Dit kan direct herleid worden naar de factor 'communicatie'. De factoren 'geloofwaardigheid', 'kwaliteit' en 'resultaten' zijn sterk afhankelijk van de kwaliteit van invoer van gegevens. Bij een beperkte relevantie zal men sneller genoeg nemen met minder kwaliteit. Wanneer de benchmark een hogere relevantie heeft, zal de kwaliteit van de input en de hiermee samenhangende nauwkeurigheid en herkenbaarheid vanzelfsprekend groter zijn. De factor communicatie speelt hierin ook een rol. Binnen WML wordt beperkt aan een selecte groep medewerkers over de benchmark gecommuniceerd. Hierdoor zal de relevantie en invloed van de benchmark niet toenemen. De beperkte communicatie zorgt er uiteindelijk voor dat de huidige situatie, d.w.z. de gelimiteerde evaluatie-invloed, kan blijven continueren.

De invloed van de benchmark bij Brabant Water wordt vooral verklaard door de factoren 'relevantie', 'geloofwaardigheid' en 'communicatie'. Alle belanghebbenden bij Brabant Water betitelen de benchmark als zeer geloofwaardig. Ook hier zijn door belanghebbenden vraagtekens geplaatst op detailniveau en is de kwaliteit niet voldoende. Deze vraagtekens hebben echter niet betrekking op de eigen informatie maar meer op de opgevoerde datasets van andere bedrijven. In lijn met WML heeft ook dit gevolgen voor de resultaten en de geloofwaardigheid van de benchmark. De benchmark wordt daarnaast als zeer relevant gezien. De doelstelling om in de top 3 te staan voor wat betreft alle benchmarkprocessen is één op één gekoppeld aan de benchmarkresultaten. Dit is een strategische doelstelling die door alle belanghebbenden gedragen en gecommuniceerd wordt. De benchmark wordt intern vooral ingezet als verbeterinstrument. Medewerkers worden uitgedaagd om beter en efficiënter te werken waarbij de

benchmark als leidraad wordt gebruikt. De communicatie betreffende de benchmark beperkt zich dan ook niet tot één maal per jaar een presentatie. Elk kwartaal worden de behaalde resultaten doorvertaald naar de benchmark en wordt een tussenstand gerapporteerd op afdelingsniveau. Afdelingen concurreren onderling met elkaar maar leren ook van elkaar. In de communicatie wordt de benchmark vooral getypeerd als verbeterinstrument in plaats van vergelijking- of afrekeninstrument. Hierdoor verklaart ook de factor 'communicatie' een belangrijk aspect van de evaluatie-invloed van de benchmark bij Brabant Water.

Deelvraag 10. Op welke manier verklaren de factoren in het kader van de besluitvorming- en beleidsetting de invloed van de Benchmark in het geval van WML en Brabant Water?

De invloed van de benchmark bij WML wordt vooral verklaard door de factoren 'besluitvormingskenmerken', 'politiek klimaat', 'persoonlijke kenmerken' en 'commitment'. De besluitvorming binnen WML is zo divers beschreven dat niet gesproken kan worden van een eenduidig of gestructureerd proces. Door een lager draagvlak voor besluiten is de mogelijke invloed van de benchmark als gevolg van directiebesluiten laag. De benchmark sluit echter wel goed aan bij het interne politieke klimaat (kostenbeheersing). De invloed van de benchmark bij WML komt vooral tot stand door individuen die de benchmark begrijpen en waarderen. Ze zien het potentieel van de benchmark en gebruiken deze voor het verbeteren van hun proces. De invloed vindt vooral plaats op afdelingsniveau en niet op bedrijfsniveau en is incidenteel van aard in plaats van structureel.

De invloed van de benchmark bij Brabant Water wordt vooral verklaard door de factoren 'besluitvormingskenmerken', 'politiek klimaat', 'persoonlijke kenmerken' en 'commitment'. De directie gelooft in de benchmark en heeft de benchmark tot een belangrijke doelstelling (top 3) gemaakt. Dit is in grote lijnen het gevolg van het commitment van enkele personen binnen Brabant Water naar de benchmark toe. Niet alleen committeert Brabant Water zich intern aan de benchmark, maar ook extern worden de doelstellingen gecommuniceerd en vastgelegd. Hoewel het initiatief en de besluitvorming met name bij de directie ligt, wordt de benchmark toch enorm beleefd binnen de organisatie. De besluitvormingskenmerken hebben een belangrijke rol in dit proces. De collectieve invloed bij Brabant Water wordt met name veroorzaakt door de top 3 doelstelling. Veel besluiten die worden genomen zijn gebaseerd op deze doelstelling.

Deelvraag 11. Op welke manier kan het verschil in invloed van de evaluatie verklaard worden door de evaluatiesettingfactoren en/of de besluitvorming- en beleidsettingfactoren?

Het verschil van invloed van de benchmark bij WML en Brabant Water wordt dus in grote mate verklaard door de factoren 'relevantie' en 'commitment' en in mindere mate door de factor 'besluitvormingskenmerken'. De relevantie van de evaluatie kan weer herleid worden naar de factor 'commitment'. De directeur van Brabant Water geeft aan dat hij de benchmark ontzettend belangrijk vindt en dat hij er voor zorgt dat zijn directieleden dit ook zo vinden. Hij is duidelijk zeer gecommitteerd aan de benchmark en verwacht dit ook bij anderen binnen Brabant Water. Op basis hiervan kan worden geconcludeerd dat in dit onderzoek de factoren behorende bij de besluitvorming- en beleidsetting van groter belang zijn voor de evaluatie-invloed dan de factoren behorende bij de evaluatiesetting.

5.2 Discussie naar aanleiding van het onderzoek

5.2.1 Analyseren onderzoeksresultaten ten opzichte van theoretisch kader

Om de onderzoeksresultaten van deze gevalsstudie te analyseren ten opzichte van het theoretisch kader van dit onderzoek, zijn de onderzoeksresultaten van de twee deelnemende bedrijven op individueel niveau afgezet tegen het theoretisch kader.

De resultaten van deze analyse laten een aantal grote overeenkomsten zien, maar ook een aantal duidelijke verschillen. Bij beide organisaties zijn alle drie de niveaus van invloed aanwezig. De collectieve invloed is echter bij Brabant Water beduidend sterker aanwezig dan bij WML. Ook spelen zowel de factoren behorende tot de evaluatiesetting als de factoren behorende tot de besluitvorming- en beleidsetting een rol van betekenis betreffende de aanwezigheid van evaluatie-invloed.

In het theoretisch kader is ook het motiverende veranderproces opgenomen. Bij zowel Brabant Water als WML speelt dit veranderproces geen enkele rol van betekenis. Alle evaluatie-invloed wordt veroorzaakt door het gedragsmatige veranderproces. Uit het onderzoek van Bober Bartlett kwam duidelijk naar voren dat de evaluatiesettingfactoren belangrijker waren dan de besluitvorming- en beleidsetting factoren. Uit dit scriptieonderzoek blijkt dat de factoren behorende bij de besluitvorming- en beleidsetting van groter belang zijn. Verder blijkt ook uit de onderzoeksresultaten dat, conform het theoretisch kader, de evaluatie-invloed beweegt tussen de verschillende niveaus van invloed. De evaluatie-invloed bij Brabant Water is veel dynamischer, er vinden dus veel bewegingen plaats tussen de verschillende niveaus. Bij WML is dit duidelijk minder.

5.2.2 Bijdragen onderzoeksresultaten aan wetenschappelijke kennis en literatuur

De bijdrage van de onderzoeksresultaten aan de wetenschappelijke kennis en literatuur is tweeledig. Ten eerste blijkt uit de onderzoeksresultaten dat alle drie de niveaus van invloed die door Mark en Henry (Mark & Henry, 2003) zijn geïdentificeerd, in de praktijk bij beide organisaties ook daadwerkelijk voor te komen. Zowel individuele, interpersoonlijke en collectieve invloed zijn in de praktijk aangetroffen. Doordat deze niveaus in de praktijk zijn aangetroffen, zijn ze niet meer alleen gebaseerd op een theoretische analyse. Hierdoor is de betrouwbaarheid van door Mark en Henry (Mark & Henry, 2003) geïdentificeerde niveaus van evaluatie-invloed vergroot.

Bovendien blijkt uit de onderzoeksresultaten dat één van de vier door Mark en Henry (Mark & Henry, 2004) geïdentificeerde typen verandermechanismen is aangetroffen in de praktijk. In dit onderzoek zijn twee van de vier verandermechanismen onderzocht (motivationale en gedragsmatig), waarbij het gedragsmatige verandermechanisme is aangetroffen in de praktijk. Hierdoor is de betrouwbaarheid van één type, van de door Mark en Henry (2004) geïdentificeerd verandermechanisme, vergroot.

Ten tweede blijkt uit de onderzoeksresultaten dat elf van de twaalf factoren die door Cousins en Leithwood (Cousins & Leithwood, 1986) zijn geïdentificeerd - en die de benutting van evaluaties kunnen beïnvloeden - ook een rol van betekenis spelen met betrekking tot de evaluatie-invloed. Hiermee wordt de bestaande theorie van Cousins en Leithwood (Cousins & Leithwood, 1986) uitgebreid. Zij beperkten zich tot de rol die deze factoren spelen bij de benutting van evaluaties. Eenzelfde factoren spelen dus ook een rol van betekenis bij het verkrijgen van evaluatie-invloed.

In tegenstelling tot het onderzoek van Bober en Bartlett (Bober & Bartlett, 2004) zijn niet de factoren die behoren tot de evaluatiesetting het meest van belang, maar de factoren die behoren tot de besluitvorming- en beleidsetting. In het onderzoek van Bober en Bartlett (Bober & Bartlett, 2004) werd de invloed van factoren onderzocht op de benutting van evaluatie. Dit onderzoek heeft gekeken naar de rol van dezelfde factoren op evaluatie-invloed. In dit onderzoek is gebleken dat de factoren die behoren bij de besluitvorming- en beleidsetting van groter belang zijn voor het hebben van evaluatie-invloed dan de factoren die behoren tot de evaluatiesetting. Dit zou kunnen betekenen dat de conclusie van Bober en Bartlett (Bober & Bartlett, 2004) niet van toepassing kan zijn op evaluatie-invloed en zich zou moeten beperken tot de benutting van evaluatie. Op basis van deze conclusie wordt de bestaande theorie over factoren die een rol van betekenis spelen bij de benutting van evaluaties uitgebreid.

5.2.3 Betrouwbaarheid en validiteit van het onderzoek

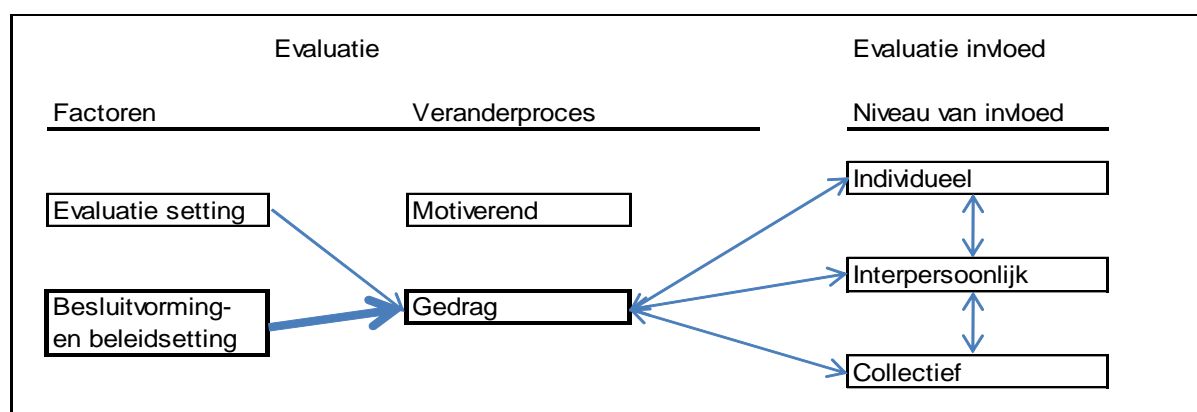
Voor de betrouwbaarheid van het onderzoek is het van belang dat, wanneer een andere onderzoeker dit onderzoek nogmaals zou uitvoeren, hij of zij tot dezelfde resultaten en conclusies komt (Yin, 2009). In de voorgaande hoofdstukken is reeds uitvoerig de keuze voor gehanteerde procedures en onderzoeks-methoden beschreven en de wijze waarop deze zijn toegepast in dit onderzoek.

Volgens Yin (Yin, 2009) kan het ontwikkelen van een database in grote mate bijdragen aan de betrouwbaarheid van een onderzoek. Om de betrouwbaarheid te vergroten is dan ook een database ontwikkeld. Hierin zijn niet alleen de interviewvragen en hun relatie tot het theoretisch kader opgenomen, maar ook alle uitkomsten van de gevoerde interviews (op basis van goedkeuring door de belanghebbenden met betrekking tot de uitkomsten van de interviews). Tevens zijn verschillende publieke documenten opgenomen die, indien mogelijk, de aanwezigheid van evaluatie-invloed kunnen aantonen. Door de uitgebreide beschrijving van de keuze en toepassing van procedures en onderzoeksmethoden en de ontwikkelde database, kan een andere onderzoeker het onderzoek herhalen en tot dezelfde resultaten en conclusies komen, hetgeen het onderzoek en de onderzoeksresultaten betrouwbaar maakt.

Volgens Yin (Yin, 2009) is ook de interne validiteit een belangrijk aspect bij een verklarend onderzoek. Een van de methoden om de interne validiteit van een onderzoek te vergroten, is gebruik te maken van de analysetechniek 'het bouwen van een verklaring' (Yin, 2009). Het uitgangspunt hierbij is het vormen van een theoretische verklaring waarbij de bevindingen uit de casus moeten leiden tot deze theoretische verklaring. Dit is de analysetechniek die in dit onderzoek is gebruikt om de onderzoeksgegevens te analyseren. De theoretische verklaringen van zowel de aanwezigheid van evaluatie-invloed als de rol die factoren behorende tot de evaluatiesetting of besluitvorming- en beleidsetting kunnen spelen, zijn onderzocht met de hulp van vragenlijsten (bijlage 1 en 2). Deze vragenlijsten zijn opgesteld op basis van het theoretisch kader van dit onderzoek. Door gebruik te maken van deze vragenlijsten als basis voor de interviews is een zekere mate van volledigheid en uniformiteit geborgd waarbij gericht data is vergaard. Hierdoor was het mogelijk om andere, niet relevante data buiten beschouwing te laten. Dit is feitelijk hetgeen de strategie van Yin (Yin, 2009) betreffende theoretische stellingen beoogt en waardoor de interne validiteit van dit onderzoek hoog is.

Door in dit onderzoek gebruik te maken van de strategie die gebaseerd is op theoretische stellingen, wordt ook de externe validiteit verhoogd (Yin, 2009). Met externe validiteit wordt bedoeld: de mate van generaliseerbaarheid of verplaatsbaarheid van onderzoeksconclusies naar andere personen, situaties, verschijnselen en tijdstippen dan die van het onderzoek. Bij kwalitatief onderzoek wordt generalisatie niet zozeer nagestreefd door gebruik te maken van een statistisch representatieve steekproef, maar wordt de onderzoeksgroep juist doelgericht samengesteld.

Voor het vaststellen van de externe validiteit is het van belang of de onderzoeksresultaten generaliseerbaar zijn buiten de gevalsstudie (Yin, 2009). De resultaten van dit onderzoek beperken zich tot WML en Brabant Water, waardoor een statische generalisatie naar andere drinkwaterbedrijven binnen Nederland of andere organisaties buiten de drinkwatersector niet mogelijk is. Een theoretische generalisatie is echter wel mogelijk. Hierbij worden de resultaten vertaald naar het theoretisch kader en het conceptueel model (Yin, 2009). Zowel de keuze voor de analyse-eenheden als de keuze voor de waarnemingseenheden is met zorg uitgevoerd en vastgelegd. Uit de onderzoeksresultaten blijkt duidelijk dat bij beide organisaties de factoren die behoren tot de categorie 'besluitvorming- en beleidsetting' meer invloed hebben gehad op de aanwezigheid van evaluatie-invloed dan de factoren die behoren tot de categorie 'evaluatie-setting'. Tevens blijkt dat de motiverende verandermechanismen bij beide organisaties geen rol van betekenis hebben gespeeld voor de evaluatie-invloed. Dit wordt in de onderliggende figuur duidelijk weergegeven.



Figuur 4 Theoretische generalisatie van onderzoeksresultaten naar conceptueel model

Door de theoretische generalisatie van de onderzoeksresultaten en de doordachte keuze met betrekking tot de analyse-eenheden en waarnemingseenheden, is de externe validiteit van dit onderzoek hoog.

5.3 Aanbevelingen naar aanleiding van het onderzoek

5.3.1 Praktische implicaties volgend uit het onderzoek

Een praktische implicatie die volgt uit dit onderzoek is dat de factoren 'commitment' en 'besluitvormings-kenmerken' die behoren tot de categorie 'besluitvorming- en beleidsetting' en de factor 'relevantie' die behoort tot de categorie 'evaluatie-setting' van belang zijn voor evaluatie-invloed. Bij de benutting van evaluatie ligt, op basis van onderzoeksresultaten Bober en Bartlett (Bober & Bartlett, 2004), de nadruk

met name op factoren die behoren tot de categorie 'evaluatie-setting'. Om een evaluatie effectief structurele invloed te kunnen laten uitoefenen binnen een organisatie zal men ook aandacht moeten besteden aan het commitment en de besluitvormingskenmerken van de organisatie. Al is de kwaliteit van de evaluatie nog zo goed, zonder voldoende draagvlak en structuur binnen de deelnemende organisaties zal dit niet leiden tot de gewenste structurele evaluatie-invloed.

Een andere implicatie die volgt uit dit onderzoek heeft betrekking op de onderzochte verandermechanismen. Bij beide organisaties is de evaluatie-invloed vooral tot stand gekomen als gevolg van de gedragsmatige verandermechanismen. De motivationele verandermechanismen spelen nauwelijks een rol van betekenis. Indien de gedragsverandering laag is, of lager dan gewenst kan men, om de evaluatie-invloed te vergroten, de aandacht wellicht focussen op datgene wat belanghebbenden triggerd om hun gedrag te veranderen.

Een andere implicatie die volgt uit dit onderzoek is het duidelijke verband tussen de factoren 'communicatie' en 'relevantie'. Uit de onderzoeksresultaten blijkt dat frequent communiceren betreffende de opzet, inhoud en resultaten van de evaluatie, positief werkt op de relevantie van de evaluatie. Een evaluatie kan dus, door regelmatig en doelgericht te communiceren, eenvoudig relevanter gemaakt worden binnen een organisatie.

5.3.2 Aanleiding voor vervolgonderzoek

De resultaten en beperkingen van dit onderzoek geven aanleiding voor meerdere vervolgonderzoeken.

Ten eerste, alle belanghebbenden binnen dit onderzoek hebben ofwel een rol gehad binnen het evaluatieproces zelf ofwel hebben ze toegang tot de volledige evaluatieresultaten. Tevens zijn bij beide organisaties belanghebbenden geselecteerd met een gelijksoortige functie. De belanghebbenden vormen een sterk homogene groep waarbinnen weinig differentiatie is. In een vervolgonderzoek kunnen ook andere medewerkers worden betrokken die niet voldoen aan bovenstaande profielschets. Evaluatie-invloed hoeft zich niet te beperken tot deze selecte groep belanghebbenden maar kan bedrijfsbreed plaatsvinden en gemeten worden.

Dit onderzoek heeft zich beperkt tot de organisaties WML en Brabant Water. In een vervolgonderzoek kunnen ook andere, of alle, Nederlandse drinkwaterbedrijven worden betrokken. Op deze manier kan onderzocht worden of deze onderzoeksresultaten geldend zijn voor de totale drinkwatersector.

De derde aanleiding voor vervolgonderzoek heeft te maken met een van de beperking van dit onderzoek. Dit onderzoek heeft zich beperkt tot de motivationele en gedragsmatige verandermechanismen. In een vervolgonderzoek kunnen ook de generieke en cognitieve en affectieve verandermechanismen worden opgenomen. Hierdoor ontstaat een verbreding van het onderzoek en wordt evaluatie-invloed in de volledige breedte onderzocht.

Dit geldt ook voor de onderzochte factoren. Dit onderzoek heeft zich beperkt tot de genoemde twaalf factoren van Cousins en Leithwood (Cousins & Leithwood, 1986). De afgelopen decennia hebben veel onderzoekers zich gefocust op factoren die de benutting van evaluaties beïnvloeden. Een recent

voorbeeld is de literatuurstudie van Johnson et Al. (Johnson, Greenesid, Toal, King, Lawrenz, & Volkov, 2009) naar de benutting van evaluaties. Zij constateerden dat de literatuurstudies van Shulha en Cousins (1997) en Cousins en Leithwood (1986) niet volledig waren. Niet alle bestaande onderzoeken waren opgenomen en ook het aspect 'kwaliteit van het onderzoek' was niet opgenomen. In een vervolgonderzoek kunnen deze nieuwe factoren worden betrokken en meegenomen als object van onderzoek (bijv. de bekwaamheid van de evaluator of stakeholder betrokkenheid (Johnson, Greenesid, Toal, King, Lawrenz, & Volkov, 2009). De factoren die evaluatie-invloed kunnen beïnvloeden laten zich niet beperken tot de eerder genoemde twaalf factoren. Door recent geïdentificeerde factoren en categorieën mee te nemen in een vervolgonderzoek, sluiten de onderzoeksresultaten aan bij de tijdsgeest. Immers, als gevolg van technologische of organisatorische ontwikkelingen kan het belang van factoren wijzigen.

6. LITERATUUR

6.1 Wetenschappelijke Literatuur

Alkin, M. (1985). *A guide for Evaluation Decision Makers*. Beverly Hills, CA: SAGE.

Alkin, M. (1990). *Debates on Evaluation*. Thousand Oaks, CA: Sage.

Alkin, M. (2003). Introduction (to Section on Evaluation Use). In T. Kelligan, & D. Stufflebeam, *International Handbook of Educational Evaluation* (pp. 189-196). Dordrecht: Kluwer Academic Press.

Alkin, M., Daillak, R., & White, P. (1979). *Using Evaluations: Does it make a difference?* Thousand Oaks, CA: Sage.

Amo, C., & Cousins, J. (2007). Going through the process: An examination of the operationalization of process use in empirical research on evaluation. *New Directions for Evaluation* , 5-26.

Ayers, T. (1987). Stakeholders as partners in evaluation: A stakeholder-Collaborative approach. *Evaluation and Program Planning* , 263-271.

Balthasar, A. (2006). *The effects of institutional Design on the Utilization of Evaluation: Evidenced Using Qualitative Comparative Analysis (QCA)*. Beverly Hills, CA: Sage.

Beyer, J., & Trice, H. (1982). The utilization proces: A conceptual framework and synthesis of empirical findings. *Administrative Science Quarterly* , 591-622.

Blokland, M., Braadbaart, O., & Schwartz, K. (1999). *Private Business, Public owners. Government Shareholdings in Water Enterprises*. The Hague: VROM.

Bober, F., & Bartlett, R. (2004). The Utilization of Training Program Evaluation in Corporate Universities. *Human Resource Development Quarterly* , 2004 (15), 363-383.

Christie, C. (2007). Reported influence of evaluation data on decision makers' Actions: An empirical examination. *American Journal of Evaluation* , 8-25.

Cousins, J. (2003). Utilization Effects of Participatory Evaluation. In T. Kelligan, & D. Stufflebeam, *International Handbook of Educational Evaluation* (pp. 245-266). Dordrecht: Kluwer Academic Press.

Cousins, J., & Leithwood, K. (1986). Current empirical research in evaluation utilization. *Review of Educational Research* , 331-364.

Cronbach, L., Ambron, S., Dornbusch, S., Hess, R., Homik, R., & Phillips, D. (1980). *Toward reform of program evaluation*. San Francisco: Jossey-Bass.

Cummings, R. (2002). Rethinking Evaluation Use. *Paper presented at the 2002 Australasian Evaluation Society International Conference* (pp. 1-7). Wollongong, Australia: Murdoch University.

- Dattakumar, R., & Jagadeesh, R. (2003). A review of literature on benchmarking. *Benchmarking: An International Journal* , 176-209.
- Dorsch, J., & Yasin, M. (1998). A framework for benchmarking in the public sector. *International Journal of Public Sector Management* , 91-115.
- Fetterman, D. (2003). A proces use distinction and a theory. *New Directions for Evaluation* , 47-52.
- Greene, J. (1988b). Stakeholder Participation and Utilization in Program Evaluation. *Evaluation Review* , 91-116.
- Harnar, M., & Preskill, H. (2007). Evaluators' descriptions of process use: An exploratory study. *New Directions for Evaluation* , 47-52.
- Henry, G. (2000). Why not use? *New Directions for Evaluation* , 85-98.
- Henry, G., & Rog, D. (1998). A realist theory and analysis of utilization. *New Directions for Evaluation* , 27-44.
- Huberman, M., & Gather Thurler, M. (1991). *De la recherche á la pratique: Eléments de base*. Bern: Peter Lang.
- Johnson, K., Greenesid, L., Toal, S., King, J., Lawrenz, F., & Volkov, B. (2009). Research on Evaluation Use: A Review of the Emperical Literature From 1986 to 2005. *American Journal of Evaluation* , 377-410.
- Johnson, R. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning* , 93-110.
- King, J., & Thompson, B. (1983). Research on school use program evaluation: A literature review and research agenda. *Studies in Educational Evaluation* , 5-21.
- Kirkhart, K. (2000). Reconceptualizing Evaluation Use: An Integrated Theory of Influence. In V. C. Preskill, *The Expanding Scope of Evaluation Use*. *New Directions for Evaluation* (pp. 5-23). San Francisco: Jossey-Bass.
- Lawrenz, F., Gullickson, A., & Toal, S. (2007). Dissimination: Handmaiden to Evaluation Use. *American Journal of Evaluation* , 275-289.
- Lawrenz, F., Huffman, D., & McGinnis, J. (2007). Multilevel evaluation proces use in large scale multisite evaluation. *New Directions for Evaluation* , 75-85.
- Leviton, L., & Hughes, E. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review* , 525-548.
- Mark, M., & Henry, G. (2003). Beyond use: Understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation* , 293-314.

Mark, M., & Henry, G. (2004). The Mechanisms and Outcomes of Evaluation Influence. *Evaluation* (10), 35-57.

Melkers, J., & Willoughby, K. (2001). Budgeters Views of State Performance Budgeting Systems: Distinctions across Branches. *Public Administration Review* , 54-64.

Patton, M. (1998). Discovering process use. *Evaluation* , 225-233.

Patton, M. (2007). Process Use as usefulness. *New Directions for Evaluation* , 99-112.

Patton, M. (1997). *Utilization-Focused Evaluation: The New Century Text*. Thousand Oaks, CA: Sage.

Patton, M., Grimes, P., Guthrie, K., Brennan, N., French, B., & Blyth, D. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. Weiss, *Using social research in public policy making* (pp. 141-164). Lexington, MA: D.C. Heath.

Preskill, H., & Caracelli, C. (1997). Current and developing conceptions of use: Evaluation use TIG survey results. *American Journal of Evaluation* , 209-225.

Preskill, H., & Torres, R. (1996). From evaluation to evaluative inquiry for organizational learning. *From evaluation to evaluative inquiry for organizational learning*. Atlanta: American Evaluation Association.

Rothman, J. (1980). *Using research in organizations: A guide to successful application*. Beverly Hills, CA: Sage.

Shadish, W., Cook, T., & Leviton, L. (1991). *Foundations of program evaluations: Theories of practise*. Newbury Parc, CA: Sage.

Shulha, L., & Cousins, J. (1997). Evaluation use: Theory, research, and practise since 1986. *American Journal of Evaluation* , 195-208.

Siegel, K., & Tuckel, P. (1985). The utilization on evaluation research. *Evaluation Review* , 307-328.

The Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards*. Thousand Oaks, CA: Sage.

Weiss, C. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation* , 21-33.

Weiss, C., & Bucuvalas, M. (1977). The challenge of social research to decision making. In C. Weiss, *Using social research in public policy making*. Lexington, MA: D.C. Heath.

Whitmore, E. (1991). Evaluation and Empowerment: It's the Process That counts. Cornell: Cornell University Empowerment Project.

Wubben, E., & Hulsink, W. (2003). A waterloo of utility liberalization? How great deregulation expectations were dashed by the Dutch water industry in the 1990's. In E. Wubben, & W. Hulsink, *On Creating Competition and Strategic Restructuring. Regular Reform in Public Utilities* (pp. 185-218). Cheltenham: Edward Elgar.

Yin, R. (2009). *Case Study Research: Design and methods*. Thousand Oaks, CA: Sage.

6.2 Documenten

Afdeling Bedrijfscontrol (2011). *Presentatie RvC MN 2012-2016 def 1.0*. WML, Maastricht.

Bierens, J., Coppens, E. (2007). *Analyse benchmarkproces Bewaken waterkwaliteit PO-BWK*. Brabant Water, Den Bosch.

Eestermans, C., Smits, T., Martens, W., de Schepper, J. (2011). *Krachtig op koers 2016 , de regio's op weg naar 2016* . Brabant Water, Den Bosch.

Peters, E., Rouleaux, J., Hoogmans, P. (2009). *Bevindingen en aanbevelingen n.a.v. bedrijfsbezoek Brabant Water in relatie tot de benchmark 2009*. WML, Maastricht.

BIJLAGE 1 VRAGENLIJST FACTOREN INTERVIEW BELANGHEBBENDEN

1. Evaluatie setting

1.1 De kwaliteit van de evaluatie

- 1.1.1 Voldoet de Benchmark aan uw kwaliteitseisen?
- 1.1.2 Welke onderdelen van de Benchmark voldoen aan uw kwaliteitsnorm en welke niet?
- 1.1.3 Hoe belangrijk is de kwaliteit van de Benchmark en waarom?

1.2 De geloofwaardigheid van de evaluatie

- 1.2.1 Hoe belangrijk is de geloofwaardigheid van de Benchmark voor u en waarom?
- 1.2.2 Hoe geloofwaardig is de Benchmark voor uw organisatie?

1.3 De relevantie van de Benchmark

- 1.3.1 Is de Benchmark belangrijk binnen uw organisatie?
- 1.3.2 Hoe relevant is de Benchmark voor u zelf en waarom?

1.4 De communicatiekwaliteit van de Benchmark

- 1.4.1 Wordt er over de Benchmark voldoende gecommuniceerd binnen uw organisatie?
- 1.4.2 Is de kwaliteit van deze communicatie voldoende en in lijn met uw verwachtingen en waarom?
- 1.4.3 Hoe belangrijk is de communicatiekwaliteit voor u en waarom?

1.5 De resultaten van de evaluatie

- 1.5.1 Geeft de Benchmark de juiste resultaten weer?
- 1.5.2 Zijn de resultaten van de Benchmark van belang voor uw organisatie en waarom?

1.6 De timing van de Benchmark

- 1.6.1 Wordt de Benchmark op tijd opgestart en worden de resultaten op tijd opgeleverd?
- 1.6.2 Hoe belangrijk is de timing van de Benchmark voor u en waarom? Verschilt uw belang met die van de organisatie?

2. Besluitvorming- en beleidsetting

2.1 De informatienoodzaak

- 2.1.1 Is de informatie die de Benchmark bevat noodzakelijk voor uw organisatie en waarom?
- 2.1.2 Wat is het belang van de informatienoodzaak voor de Benchmark?
- 2.1.3 Is er behoefte aan de informatie van de Benchmark en welke informatie is dit of zou dat moeten zijn?

2.2 De besluitvormingskenmerken van de organisatie

- 2.2.1 Wat zijn volgens u de besluitvormingskenmerken van de organisatie?
- 2.2.2 Zijn de besluitvormingskenmerken van de organisatie van belang voor de Benchmark?

2.3 Het politieke klimaat

- 2.3.1 Speelt de politiek een rol bij de Benchmark? Wie heeft volgens u belang bij de Benchmark en wie heeft er belang bij om de resultaten van de Benchmark stil te houden en waarom? Speelt dit laatste een rol van betekenis bij de Benchmark?
- 2.3.2 Sluit de Benchmark aan bij de wensen van zowel het interne- als externe politieke klimaat?
- 2.3.3 Wat is volgens u het politieke klimaat waarin uw organisatie opereert (intern als extern)?

2.4 Concurrerende informatie

- 2.4.1 Is de informatie die de Benchmark bevat uniek en waardevol voor uw organisatie of bestaan er andere bronnen van informatie die dezelfde gegevens bevatten en even betrouwbaar zijn?
- 2.4.2 Wat is het belang van de concurrerende informatie voor de Benchmark?

2.5 Persoonlijke kenmerken van de gebruiker

- 2.5.1 Hoe zou u u zelf typeren in termen van persoonlijke kenmerken (open, leergierig)? Hoeveel jaren ervaring heeft u met de Benchmark?
- 2.5.2 Wat is uw functie, opleidingsniveau?
- 2.5.3 Wat is volgens u het belang van de persoonlijke kenmerken voor de Benchmark en waarom?

2.6 Commitment naar de evaluatie toe

- 2.6.1 Hoe staat u zelf, persoonlijk tegenover de Benchmark (positief, negatief)? Bent u bereid de resultaten van de Benchmark te gebruiken en waarom?
- 2.6.2 Wat is het belang van het commitment van zowel de organisatie als individuen voor de Benchmark?

BIJLAGE 2 VRAGENLIJST INVLOED INTERVIEW BELANGHEBBENDEN

Invloed

1. Individueel

- 1.1 Heeft u als gevolg van de Benchmark uw houding veranderd ten aanzien van specifieke onderwerpen?
- 1.2 Heeft u als gevolg van de Benchmark nieuwe vaardigheden verworven of bent u bestaande vaardigheden anders gaan gebruiken / inzetten?
- 1.3 Heeft u als gevolg van de Benchmark uw gedrag veranderd ten aanzien van specifieke situaties?

2. Interpersoonlijk

- 2.1 Heeft u met behulp van de Benchmark andere kunnen aanzetten tot verandering (ook organisatorische verandering)?
- 2.2 Heeft u als gevolg van de Benchmark uw sociale normen gewijzigd?

3. Collectief

- 3.1 Is het beleid van organisatie veranderd als gevolg van de Benchmark?
- 3.2 Is het beleid van uw organisatie gewijzigd/veranderd als gevolg van de resultaten van de Benchmark buiten de Nederlandse drinkwatersector (ongewild)?

BIJLAGE 3 INTERVIEWVERSLAGEN BELANGHEBBENDEN

In de digitale versie van dit onderzoek zijn de interviewverslagen wegens geheimhouding niet in de bijlagen opgenomen. De interviewverslagen zijn wel ter beschikking gesteld aan de faculteit managementwetenschappen van de OU.